NESTED LOGIT OR RANDOM COEFFICIENTS LOGIT? A COMPARISON OF ALTERNATIVE DISCRETE CHOICE MODELS OF PRODUCT DIFFERENTIATION

Laura Grigolon and Frank Verboven*

Abstract—We propose a random coefficients nested logit (RCNL) model to compare the tractable nested logit (NL) model with the more complex random coefficients logit (RC) model. After a simulation study, we use data on the European automobile market. Both the NL and RC models are rejected against the RCNL model. The RC model results in different substitution patterns and a wider market definition than the NL and RCNL models. Nevertheless, the predicted price effects from mergers are robust across models. Our findings stress the importance of accounting for discrete sources of market segmentation not captured by continuous product characteristics.

I. Introduction

DISCRETE choice models of product differentiation have gained considerable importance in empirical work. Because they treat products as bundles of characteristics, they offer the possibility of uncovering rich substitution patterns with a limited number of parameters. Berry (1994) developed a framework to estimate a class of discrete choice models with unobserved consumer heterogeneity based on aggregate sales data. His framework includes the random coefficient logit model of Berry, Levinsohn, and Pakes (1995, hereafter BLP), the nested logit model (with special random coefficients on discrete product characteristics), and the logit model (without consumer heterogeneity).

The logit and nested logit models have been popular because of their computational simplicity, since they can be transformed to simple linear regressions of market shares on product characteristics. At the same time, they have long been criticized because they yield substitution patterns that are too restrictive. The logit model assumes that consumer preferences are uncorrelated across all products, implying symmetric cross-price elasticities. The nested logit model allows preferences to be correlated across products within the same group or "nest." It thus entails a special kind of random coefficients on group dummy variables (Cardell, 1997). It allows products of the same group to be closer substitutes than products of different groups, but the aggregate substitution patterns remain restrictive: cross-price elasticities within the same group are still symmetric, and substitution outside a group is symmetric to all other groups. In contrast, BLP's random coefficients logit model incorporates random coefficients for continuously measured product characteristics (and at the same time still allows random coefficients on group dummy variables with other distributional assumptions than those of the nested logit model). This creates potentially more flexible substitution patterns, where products tend to be closer substitutes as they have more similar continuous characteristics. However, the random coefficients model is computationally more demanding, and several recent papers have studied a variety of problems relating to its numerical performance (see Knittel & Metaxoglou, forthcoming; Dubé, Fox, & Su, 2012; Judd & Skrainka, 2011).

Against this background, a particularly timely issue is to assess whether and when the popular logit and nested logit models can be used as reasonable alternatives to the computationally more demanding random coefficients logit model. In this paper we provide a systematic comparison of these demand models, and as an illustration, we assess how they perform in competition policy analysis. To accomplish this, we start from a random coefficients nested logit model (RCNL), which combines the random coefficients logit (RC) and nested logit (NL) models. The random coefficients on the continuous characteristics can take any distributional form, as in the general RC framework. In contrast, the random coefficients for the discrete characteristics take the special distributional assumptions of the NL model. This simplifies the computational burden and enables us to assess the relative importance of both sources of consumer heterogeneity.[1]

To motivate our empirical analysis, we begin with two groups of Monte Carlo experiments. First, we consider an RC model with a normally distributed random coefficient for a group dummy variable. For a wide variety of designs, we find that the true RC model and a misspecified NL model result in similar estimated own- and cross-price elasticities. Hence, the specific distributional assumptions of the RC and NL models regarding the valuation for the group dummy variable do not matter much in this simple setup.[2] Second, we consider RCNL models with both a normally distributed random coefficient for a continuous characteristic and a nesting parameter for a group dummy variable. For the wide range of considered designs, we find that both the RC and the NLs model are reasonable approximations, with stronger substitution within

Received for publication September 17, 2011. Revision accepted for publication August 14, 2013. Editor: Philippe Aghion.

* Grigolon: McMaster University; Verboven: University of Leuven and CEPR.

We are grateful to three referees and to Geert Dhaene, Tobias Klein, Pasquale Schiraldi, Øyvind Thomassen, Jo Van Biesebroeck, and Patrick Van Cayseele for their useful comments. We also thank seminar participants at TSE, and conference participants at EARIE 2010, ZEW 2010, CRESSE 2012, and RES 2012.

A supplemental appendix is available online at http://www.mitpressjournals.org/doi/suppl/10.1162/REST_a_00420.

[1] As shown by McFadden and Train (2000), any discrete choice model can be approximated by an RC model. This of course also applies to our setting. For example, it is possible to include random coefficients for the group dummy variables directly in an RC model (with other distributional assumptions than those of the nested logit). This would, however, be computationally very costly when there are many group dummy variables.

[2] This is also confirmed by a reverse set of Monte Carlo where the true data-generating process is that of an NL model and the RC is the misspecified model.

The Review of Economics and Statistics, December 2014, 96(5): 916–935

doi:10.1162/REST_a_00420

than between groups. The RC model provides a better approximation of the own-price elasticities than the NL model, while both models tend to underestimate the cross-price elasticities within a group.

These findings are confirmed in our main empirical analysis. We collected a unique data set on the automobile market for nine European countries covering around 90% of the car sales in the European Union from 1998 to 2006. The market is commonly classified in various segments (subcompact, compact, intermediate, standard, luxury, SUV, and sports), and car manufacturers typically promote their models as belonging to one of these segments. Hence, the segments may proxy for observed product characteristics such as size, engine performance, and fuel efficiency. But it is also possible that they capture intrinsically unobserved features shared by different car models. Our data set is therefore particularly interesting to compare the performance of the logit, NL, RC, and RCNL models. Consistent with earlier findings, the logit model is rejected against both the NL and RC models. More important, in the RCNL model, the nesting parameters become quantitatively smaller (consistent with the results of our Monte Carlo experiment), but they remain highly significant and economically important. Furthermore, the random coefficients relating to car size become insignificant, while the random coefficients relating to engine power and fuel efficiency remain significant. These various findings suggest that the nesting parameters may proxy for random coefficients of some of the observed continuous characteristics but also capture other unobserved dimensions of consumer preferences.

To illustrate the implications of our findings, we present own- and cross-price elasticities for the different models, and we perform policy counterfactuals common in competition policy: market definition and merger simulation. In terms of substitution patterns, the results are consistent with our Monte Carlo analysis: the RC model better approximates the own-price elasticities than the NL model, and both models somewhat underestimate the cross-price elasticities within segments (but less so for the NL model). Despite the different substitution patterns, merger simulations of two domestic mergers yield robust conclusions across different demand models: while the simple logit clearly appears inappropriate, the NL, RC, and RCNL all tend to give robust conclusions. In contrast, the conclusions for market definition are less robust: the RC suggests a wider market definition than the NL and RCNL models. This is true whether we start from "segments" or from "ten nearest substitutes" as the candidate relevant markets. We draw two implications for competition policy. First, in market definition, it is important to include sufficient sources of market segmentation in the demand model. Second, the robustness in merger simulation suggests that the simple NL model can be sufficient to obtain reliable policy conclusions, despite the different substitution patterns from the RC model.

More generally, one can draw two implications for the choice of demand model in applied work. First, the choice between the tractable NL model and the computationally more complex RC model may depend on the application. In our analysis of hypothetical domestic mergers, consumer heterogeneity regarding the cars' domestic or foreign origin is particularly relevant, and the NL model captures this reasonably well. In other applications, the most relevant aspects of consumer heterogeneity may not be captured well by nesting parameters for groups or subgroups. In these cases, it is appropriate to estimate RC models with random coefficients for the most relevant continuous characteristics.

Second, our results imply that it can be important to account for sources of market segmentation that are not captured by continuously measured product characteristics. Adding a nested logit structure to BLP's random coefficients model is a tractable way to accomplish this, since it gives closed-form expressions for integrals in the choice probabilities.[3] In principle, BLP's framework can of course also incorporate random coefficients on group dummies (and in a more flexible way). But this is more complicated because it increases the dimensionality of the integrals that need to be simulated, and in practice it often proves difficult to estimate the coefficients as precisely as in the closed-form GEV models. For example, Nevo (2001) estimates a rich demand model for the U.S. cereals market. His model includes three random coefficients for the segments (all-family, kids, and adult), but two of these are estimated rather imprecisely.

Our comparison of alternative discrete choice models is timely for several related reasons. First, a few recent papers have thoroughly studied several (often commonly known) numerical difficulties with the aggregate random coefficients model: the role of starting values and different optimization methods (Knittel & Metaxoglou, forthcoming), the accuracy of BLP's contraction mapping to invert the market share system (Dubé et al., 2012), and alternative methods of integration of the market share system (Judd & Skrainka, 2011). We draw from these findings in our own empirical analysis by cautiously considering multiple starting values, using a tight inner loop contraction mapping, and carefully approximating the market share integrals.[4]

Second, there is a large and rapidly growing empirical literature estimating aggregate discrete choice models of product differentiation, with applications in industrial organization, international trade, environmental and public economics, marketing, finance, and other areas. A complete review of the applied aggregate discrete choice literature is beyond the

[3] Instead of the nested logit structure, one may also consider other tractable models from McFadden's (1978) generalized extreme value model (GEV). Examples are Small's (1987) model of ordered alternatives, Bresnahan, Stern, and Trajtenberg's (1997) principles of differentiation model, and most recent, the flexible coefficient multinomial logit model of Davis and Schiraldi (2012).

[4] To invert the market share system, we do not consider Dubé et al.'s (2012) alternative MPEC approach because we have a large number of products and markets, implying a large number of nonlinear constraints in their constrained optimization algorithm. To approximate the market share integrals, we follow Judd and Skrainka (2011) in our Monte Carlo with low-dimensional integrals and use a large number of Halton draws in our empirical analysis with high-dimensional integrals.

scope of this introduction, so we limit attention here to early work. Much of this work has actually also looked at automobiles. Bresnahan (1981) and Feenstra and Levinsohn (1995) are important contributions preceding the seminal work of Berry (1994) and BLP. Verboven (1996) and Fershtman and Gandal (1998) are early applications of Berry's (1994) aggregate nested logit model. Nevo (2001), Petrin (2002), and Sudhir (2001) are early applications with interesting extensions of BLP's full random coefficients model. In recent years, academic work appears to focus more exclusively on the random coefficients models, whereas competition policy practitioners often use the logit and nested logit models. Our findings on the automobile market suggest that the nested logit model may be a reasonable approximation not only in competition policy but also in other applications where the market segments are the most relevant differentiating dimensions, for example, an analysis of trade liberalization. In contrast, applications on, for example, quality discrimination or environmental policy would warrant estimating BLP's random coefficients logit model, since the relevant random coefficients (engine power and fuel efficiency) are not well captured by the nesting parameters.[5]

The rest of this paper is organized as follows. Section II presents the model and conducts Monte Carlo experiments. Section III uses the data set for the European car market to estimate the logit, NL, RC, and RCNL models and the implied price elasticities. Section IV draws implications for competition policy analysis, applying market definition and merger simulation. Conclusions follow in section V.

II. The Model

A. Demand

We consider an RCNL that contains the logit, NL, and RC as special cases. As we discuss, the RCNL model can itself be seen as a random coefficients model, with special distributional assumptions for the heterogeneity on the discrete product characteristics.

There are T markets, $t = 1, \ldots, T$. In each market t, there are L_t potential consumers. Each consumer i may either choose the outside good 0 or one of the J differentiated products, $j = 0, \ldots, J$. Consumer i's conditional indirect utility for the outside good is $u_{i0t} = \bar{\varepsilon}_{i0t}$. For products $j = 1, \ldots, J$, it is

$$u_{ijt} = x_{jt}\beta_i + \xi_{jt} + \bar{\varepsilon}_{ijt}, \tag{1}$$

where x_{jt} is a $1 \times K$ vector of observed product characteristics (including price), β_i is a $K \times 1$ vector of random coefficients capturing the individual-specific valuations for the product characteristics, ξ_{jt} refers to unobserved product characteristics (to the econometrician), and $\bar{\varepsilon}_{ijt}$ is a remaining individual-specific valuation for product j.

Assume that the distributions of the individual valuations β_i and $\bar{\varepsilon}_{ijt}$ are mutually independent. The random coefficients vector, β_i, can be specified as follows. Let β be a $K \times 1$ vector of mean valuations of the characteristics, σ be a $K \times 1$ vector with standard deviations of the valuations, and ν_i be a $K \times 1$ vector with standard normal random variables. We then specify

$$\beta_i = \beta + \Sigma\nu_i, \tag{2}$$

where Σ is a $K \times K$ diagonal matrix with the standard deviations σ on the diagonal.[6]

The individual valuations for the products j, $\bar{\varepsilon}_{ijt}$ follow the distributional assumptions of the NL model, which allows valuations to be correlated across products in the same group. More specifically, suppose each product j can be assigned to one of G collectively exhaustive and mutually exclusive groups, $g = 0, \ldots, G$, where group 0 is reserved for the outside good 0. Write

$$\bar{\varepsilon}_{ijt} = \zeta_{igt} + (1 - \rho)\varepsilon_{ijt}, \tag{3}$$

where ε_{ijt} is i.i.d extreme value and ζ_{igt} has the (unique) distribution such that $\bar{\varepsilon}_{ijt}$ is extreme value. The nesting parameter ρ, with $0 \leq \rho < 1$, proxies for the degree of preference correlation between products of the same group. As ρ goes to 1, the within-group correlation of utilities goes to 1, and consumers perceive products of the same group as perfect substitutes. As ρ goes to 0, the within-group correlation goes to 0, and the model reduces to the simple logit.

Following Berry's (1994) discussion of Cardell (1997), the ζ_{igt} can be interpreted as random coefficients on group-specific dummy variables. More specifically, let d_{jgt} be a group dummy variable equal to 1 if j belongs to group g. Using equations (2) and (3) and defining the mean utility for product j, $\delta_{jt} \equiv x_{jt}\beta + \xi_{jt}$, we can write consumer i's conditional indirect utility (1) as

$$u_{ijt} = \delta_{jt} + x_{jt}\Sigma\nu_i + \sum\nolimits_{g=1}^{G} d_{jgt}\zeta_{igt} + (1 - \rho)\varepsilon_{ijt}. \tag{4}$$

Indirect utility thus contains consumer heterogeneity for the continuous characteristics x_{jt} and consumer heterogeneity for the group dummy variables d_{jgt}. The heterogeneity in the valuations for d_{jgt} follows the specific distributional assumptions of the nested logit. As an alternative, one may include d_{jgt} in the x_{jt} vector. This potentially creates more flexibility; in particular, it also allows one to incorporate correlation in the random coefficients for the continuous and discrete

[5] Wojcik (2000) also compares the NL and RC models. She claims the NL model is likely to be superior, but Berry and Pakes (2001) raise serious methdological problems with her comparison. Our approach is rather different from Wojcik's since we start from an RCNL model that enables a better comparison of the NL and RC models. Furthermore, we follow prediction excercises in the spirit of those advocated by Berry and Pakes by focusing on comparing price elasticities and other counterfactuals. Our conclusions are much more nuanced since we focus on identifying circumstances where the NL may, or may not, be a reasonable alternative.

[6] In principle, one may also specify nonzero off-diagonal elements in Σ to allow consumer valuations to be correlated across characteristics.

characteristics. This will come, however, at an increased computational cost. We can see several special cases in equation (4). If $\sigma_k = 0$ for all elements in Σ, we obtain the standard nested logit model. If $\rho = 0$, we obtain BLP's random coefficients logit model. And if all $\sigma_k = \rho = 0$, the simple logit model results.

Each consumer i in market t chooses the product j that maximizes her utility. With the above assumptions, the conditional probability that consumer i chooses product j in market t takes the nested logit form,

$$\pi_{jt}(\delta_t, \theta, \nu_i) = \frac{\exp\left(\left(\delta_{jt} + x_{jt}\Sigma\nu_i\right) / (1-\rho)\right)}{\exp\left(I_{ig} / (1-\rho)\right)} \frac{\exp I_{ig}}{\exp I_i}, \quad (5)$$

where $\theta = (\beta, \sigma, \rho)$ and I_{ig} and I_i are McFadden's (1978) "inclusive values" defined by

$$I_{ig} = (1-\rho) \ln \sum\nolimits_{k=1}^{J_g} \exp\left(\left(\delta_{kt} + x_{kt}\Sigma\nu_i\right) / (1-\rho)\right),$$
$$I_i = \ln\left(1 + \sum\nolimits_{g=1}^{G} \exp I_{ig}\right),$$

and J_g is the number of products in segment g such that $\sum_{g=1}^{G} J_g = J$. The unconditional choice probability or aggregate market share of product j in market t is

$$s_{jt}(\delta_t, \theta) = \int \pi_{jt}(\delta_t, \theta, \nu)\phi(\nu)d\nu, \quad (6)$$

which simplifies to BLP's random coefficients logit model if $\rho = 0$. The market share integral may be approximated using Monte Carlo integration as in BLP or using quadrature rules as suggested by Judd and Skrainka (2011).

To estimate the demand parameters θ, we follow Berry (1994), BLP, and the subsequent literature. We equate the observed market share vector (i.e., unit sales per product divided by the number of potential consumers L_t) to the predicted market share vector, $s_t = s_t(\delta_t, \theta)$. We solve this system for δ_t in each market t using fixed-point iteration on

$$\delta_t^{r+1} \equiv \delta_t^r + \ln(s_t) - (1-\rho)\ln(s_t(\delta_t^r)) \quad (7)$$

until convergence is reached ($\delta_t^{r+1} \approx \delta_t^r$). This is a slight modification of BLP's contraction mapping: it dampens the final term by $1-\rho$ (see also Brenkers & Verboven, 2006). In appendix A we show that this dampening is necessary to satisfy the conditions for a contraction mapping in nested logit models when ρ is sufficiently high (as opposed to logit models where dampening is not needed). Since the error term enters additively in δ_t, solving for δ_t gives a solution for the error term ξ_{jt} for each product $j = 1, \ldots, J$ in market t. We can then interact this with a set of instruments providing the moment conditions to proceed with GMM, as we discuss in more detail in section 3.

B. *Monte Carlo Experiments*

Overview. Before analyzing the car market data we consider two setups of Monte Carlo experiments. Setup 1 generates data sets according to RC models. We specify one normally distributed random coefficient for a discrete product characteristic, a dummy variable for the product's group. We focus on the performance of estimating misspecified NL models, where the nests are defined according to the same group dummy variable. For a wide range of designs, we find that the NL model approximates the substitution patterns of the RC model very well, so the different distributional assumptions do not matter much.

Setup 2 goes a step further and also includes consumer heterogeneity for a continuous product characteristic. We generate data sets according to the more general RCNL models, with one normally distributed random coefficient for a continuous product characteristic (price) and one nesting parameter for the group dummy variable. Across all our designs, we obtain the following robust conclusions. While the logit, NL, and RC models yield biased parameter estimates of the true RCNL model, only the logit model generates implausible substitution patterns. The RC model provides a better approximation of the own-price elasticities than the NL model. Furthermore, in contrast to the logit model, both the RC and the NL models yield stronger substitution within than between groups (although all models underestimate the cross-price elasticities between products of the same group).

We focus most of the discussion on substitution biases from estimating misspecified models. But we also take the opportunity to briefly comment on the numerical performance of the different models in light of the recent literature on these issues.

Setup. We conduct a large number of Monte Carlo experiments in two main setups: an RC model for a discrete characteristic and an RCNL model. For each experiment we generate 1,000 data sets. Each data set consists of $T = 50$ independent markets and $J = 25$ products per market. Each product j in each market t has an observed characteristics vector $x_{jt} = (1, x_{jt}^1, d_{jt})$: a constant, one continuous characteristic x_{jt}^1 (the price variable), and one discrete characteristic d_{jt}, a dummy variable referring to the product's group or nest (segment 0 or 1). Furthermore, each product has an unobserved product characteristic ξ_{jt}, which is drawn from a standard normal distribution and uncorrelated with x_{jt}. We thus abstract from the issue of endogeneity of some product characteristics (typically price), since we want to focus on comparing the performance of misspecified models.[7]

We draw the continuous price variable x_{jt}^1 from a lognormal distribution, which ensures positive realizations and roughly mimics the distribution of prices in empirical data sets. It will be convenient to treat the dummy variable for the

[7] For a recent discussion on the endogeneity of product characteristics, and the role of instruments in the GMM context, see Armstrong (2012) and Reynaert and Verboven (2014).

product's group d_{jt} as the realization of a latent continuous variable d_{jt}^{*}. More specifically, assume that

$$\begin{pmatrix} \ln x_{jt}^{1} \\ d_{jt}^{*} \end{pmatrix} \sim N\left(\begin{matrix}0\\0\end{matrix}, \begin{matrix} 1 & \varsigma_{xd} \\ \varsigma_{xd} & 1 \end{matrix}\right),$$

and $d_{jt} = \mathbf{1}_{\{d_{jt}^{*}>\gamma\}}$. In the Monte Carlo designs we can vary ς_{xd}, the correlation between d_{jt}^{*} and x_{jt}^{1}, to capture the extent to which the product's group is informative about the continuous characteristic. We can also vary the cutoff point γ to set the number of products allocated in groups 0 and 1. Unless stated otherwise, we set $\varsigma_{xd} = 0$ and $\gamma = 0$ as the default values.

We specify a broad range of preference parameters: the mean valuations $\beta = (\beta_0, \beta_{x^1}, \beta_d)$, the standard deviations $\sigma = (\sigma_0, \sigma_{x^1}, \sigma_d)$, and the nesting parameter ρ. We discuss the various parameter designs in more detail in the next section, where we explain the results for our two main setups.

The market shares are computed from the market share equation (6) using the generated product characteristics (x_{jt} and ξ_{jt}) and the assumed demand parameters $\theta = (\beta, \sigma, \rho)$. To approximate the market share integral (6) over the normal random variable vector ν_i, we use an accurate polynomial-based sparse grid quadrature rule as suggested by Judd and Skrainka (2011),

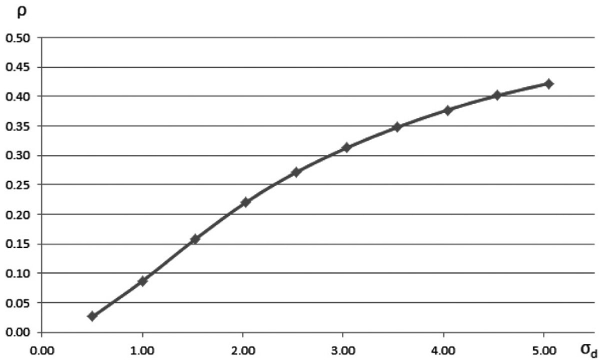
$$s_{jt}(\delta_t, \theta) \approx \sum\nolimits_{i=1}^{N} \varphi_i \frac{\exp\left(\left(\delta_{jt} + x_{jt}\Sigma\nu_i\right)/(1-\rho)\right)}{\exp\left(I_g/(1-\rho)\right)} \frac{\exp I_g}{\exp I}, \tag{8}$$

where we use nine nodes appropriately weighted by φ_i.[8] Since our focus is not on numerical integration error, we use the same set of nodes and weights to compute the market share in the DGP and the estimation.

For each design and its associated 1,000 data sets, we use GMM to estimate the correctly specified model and the other, misspecified models. In all cases, we generate the set of instruments from within the model, following Chamberlain's (1987) approach as applied in Berry et al. (1999). Given the demand parameters $\theta = (\beta, \sigma, \rho)$, this instrument vector is the expected value of $\partial\xi_{jt}(\theta)/\partial\theta'$. This includes the characteristics vector itself (x_{jt}) and nonlinear functions of the characteristics and the parameters. Finally, following the recent literature we consider multiple starting values, a state-of-the-art optimization algorithm (Knitro 8.0), and a tight inner loop contraction mapping.

Results setup 1: RC model for discrete characteristic. We specify RC models with only a random coefficient for the discrete characteristic d_{jt} and no other consumer heterogeneity. More specifically, we assume that the mean valuations are $\beta = (-1, -2, -0.5)$, the standard deviations are $\sigma = (0, 0, \sigma_d)$, and the nesting parameter is $\rho = 0$.[9] We consider ten designs for σ_d, varying over $0.5, 1, \ldots, 5$.[10] For each design and each of the 1,000 generated data sets, we estimate both the correctly specified RC model and the misspecified NL and logit models. The RC and NL models differ only in the distributional assumptions regarding consumer heterogeneity for the discrete characteristic d_{jt} (normal versus generalized extreme value). This comparison thus enables us to assess to what extent ρ takes over the role of σ_d and whether the misspecified NL model provides a good approximation for the price elasticities of the correctly specified RC model.[11]

FIGURE 1.—RELATIONSHIP BETWEEN HETEROGENEITY PARAMETERS IN THE RC AND NL MODELS



The relationship between the mean of the estimated ρ in the misspecified NL model against the mean of the estimated σ_d in the correctly specified RC model. The estimates are based on 1,000 random samples of 50 markets and 25 products. The true model is the RC model with 10 different designs: $\sigma_d = 0.5; 1 \ldots 5$.

Figure 1 plots the mean of the estimated ρ in the misspecified NL models against the mean of the estimated σ_d in the correctly specified RC models for the ten different designs. There is a clear increasing relationship between the estimated ρ and σ_d. The estimated nesting parameter ρ thus captures the omitted random coefficient σ_d. Furthermore, ρ is increasing in σ_d at a decreasing rate. Hence, ρ remains in the unit interval even for high values of σ_d, consistent with the assumptions of the NL model.

Table 1 shows more detailed estimation results for two of the ten designs: limited consumer heterogeneity for the discrete characteristic ($\sigma_d = 1$) and strong heterogeneity ($\sigma_d = 5$). The top part shows the mean and standard deviation of the main parameters as estimated from the 1,000 generated data sets. We first look at the parameters for the correctly specified RC model. For both designs, the means of β_{x^1} and σ_d are very close to the true parameters, their standard deviations are fairly small, and the distribution is approximately normal (as confirmed by Q-Q plots, not shown). This confirms that

[8] A relatively low number of nodes is needed for accurate integration: with nine nodes, the unidimensional market share integral is exact at accuracy level 17.

[9] The choice of the mean valuations β is based on obtaining a realistic setting in which the outside good market share varies between 0.10 and 0.70 under the different data-generating processes.

[10] We took the default values for the correlation between d_{jt}^{*} and x_{jt}^{1}, $\varsigma_{xd} = 0$ and for the cutoff point of belonging to group 0 or 1, $\gamma = 0$. We also experimented with different values of ς_{xd} and γ, and this gives robust conclusions.

[11] We have also conducted the reverse group of experiments: we specify an NL model and compare the estimates of the correctly specified NL model with those of a misspecified RC model. We use the same mean valuation vector and consider ten designs for the nesting parameter, $\rho = 0.10, 0.15, \ldots, 0.50$. We obtain comparable results, as reported in online appendix B.

TABLE 1.—MONTE CARLO RESULTS SETUP 1: RC MODEL FOR DISCRETE CHARACTERISTIC

	True $\sigma_d = 1$			True $\sigma_d = 5$		
Parameter	Logit	NL	RC	Logit	NL	RC
β_{x^1}	−2.00	−1.83	−2.00	−1.98	−1.16	−2.00
	(0.04)	(0.12)	(0.04)	(0.05)	(0.18)	(0.04)
σ_d	NA	NA	1.00	NA	NA	5.05
			(0.36)			(0.45)
ρ	NA	0.09	NA	NA	0.42	NA
		(0.08)			(0.12)	
Own elasticity	−2.607	−2.610	−2.607	−2.579	−2.599	−2.582
	(0.404)	(0.405)	(0.405)	(0.400)	(0.404)	(0.404)
Cross-elasticity, same segment	0.037	0.042	0.043	0.037	0.052	0.056
	(0.009)	(0.012)	(0.011)	(0.009)	(0.014)	(0.014)
Cross-elasticity, different segment	0.037	0.034	0.032	0.038	0.028	0.021
	(0.010)	(0.009)	(0.008)	(0.010)	(0.008)	(0.005)
Model selection criteria						
GMM-AIC	1	263	736	0	6	994
GMM-BIC	2	262	736	0	6	994

The table reports the empirical means and standard deviations (in parentheses) of selected parameters, the implied price elasticities for $T = 1$, and two information criteria, GMM-BIC and GMM-AIC. The estimates are based on 1,000 random samples of 50 markets and 25 products. The true model is the RC model of setup 1 with two designs: $\sigma_d = 1$ and $\sigma_d = 5$.

our estimation procedure, with a state-of-the-art optimization algorithm, analytical derivatives, numerical integration, and a tight inner loop contraction mapping, works well in practice.[12]

Next, we consider the parameters of the misspecified models. The price parameter β_{x^1} is close to the true value of -2 in the misspecified logit model. In contrast, it appears to differ significantly from the true value in the misspecified NL model, especially in the second design with strong consumer heterogeneity ($\widehat{\beta}_{x^1} = -1.16 > -2$). However, this is entirely due to the significance of the nesting parameter ρ, which rescales all utility parameters. Indeed, the rescaled price parameter $\beta_{x^1}/(1-\rho)$ is very close to the true value of -2.

What do these findings imply for the estimated price elasticities? The middle part of table 1 shows that the own-price elasticities of the logit and NL models are very close to the ones in the RC model (on average about -2.57 for all models). Furthermore, the cross-price elasticities of the NL model are also quite close to those of the RC model, with stronger substitution within than between groups (especially in the second design with $\sigma_d = 5$). In contrast, the logit model yields symmetric substitution patterns. Hence, although the NL model resulted in a biased price parameter, the nesting parameter ρ accounts for the omitted σ_d, so that the bias in the cross-price effects is very small. As expected, the NL model especially improves on the logit model in the second design with strong consumer heterogeneity ($\sigma_d = 5$).

Finally, we calculated the AIC and BIC model selection criteria, as developed in the GMM framework by Andrews (1999). Most interesting, in the design with limited consumer heterogeneity ($\sigma_d = 1$), both selection criteria have difficulties distinguishing between the misspecified NL and the true RC model: they incorrectly pick the NL model as the true data-generating process in up to 26% of the 1,000 cases. In contrast, in the design with strong consumer heterogeneity ($\sigma_d = 5$), both selection criteria correctly detect the RC model as the true model in almost all cases.

Note that we also implemented the reverse Monte Carlo experiments, where the NL model is the true data-generating process and the RC model is the misspecified model (reported in online appendix B). We can draw analogous conclusions: the RC model captures the asymmetric substitution patterns of the NL model quite well, with only minor differences because of different distributional assumptions. This finding is consistent with McFadden and Train (2000), who showed that any random utility model (so also a NL model) can be approximated by a random coefficients logit model.

To summarize, in contrast with the logit model, a misspecified NL model results in comparable asymmetric substitution patterns as the true RC model (and vice versa). In fact, it is quite difficult to formally distinguish between the NL and RC models for low levels of consumer heterogeneity. Hence, when consumer heterogeneity mainly refers to a discrete characteristic, the computationally tractable NL model may often be preferable to the RC model (unless there are strong reasons to expect a specific functional form for the distribution for consumer heterogeneity).

Results setup 2: RCNL model. We now specify RCNL models with a random coefficient for the continuous characteristic x_{jt}^1 and a nesting parameter ρ for the group dummy variable. More specifically, we assume that $\beta = (-1, -3, -2)$, $\sigma = (0, \sigma_{x^1}, 0)$ and take various values for ρ.[13] We consider eight designs according to three criteria. Firstly, set either $\sigma_{x^1} = 1.0$ and $\rho = 0.3$, or $\sigma_{x^1} = 0.5$ and $\rho = 0.5$: the first case is closer to an RC model, whereas the second case is closer to an NL model. Second, set the cutoff point

[12] Note that the nesting parameter ρ enters nonlinearly in the choice probabilities. This may be another source of multiple minima and sensitivity to starting values. However, in our GMM context, we did not find the issue to be larger than in the models without ρ.

[13] The choice of the mean valuations β is again based on obtaining a realistic setting for the outside good market share, varying between 0.20 and 0.83.

TABLE 2.—MONTE CARLO RESULTS SETUP 2: RCNL MODEL (ONE DESIGN)

Coefficients	True Parameter	Logit	NL	RC	RCNL
β_0	−1	−2.89	−0.16	−1.60	−1.00
		(0.06)	(0.20)	(0.09)	(0.16)
β_d	−2	−0.44	−5.66	−0.86	−2.01
		(0.15)	(0.46)	(0.12)	(0.30)
β_{x^1}	−3	−1.94	−0.31	−4.06	−2.99
		(0.06)	(0.13)	(0.16)	(0.27)
ρ	0.3	n/a	0.88	n/a	0.30
			(0.07)		(0.07)
σ_{x^1}	1	n/a	n/a	1.30	1.00
				(0.06)	(0.08)
Own elasticity		−2.545	−3.347	−5.205	−5.344
		(0.397)	(1.513)	(0.820)	(0.860)
Cross-elasticity, same segment		0.028	0.093	0.083	0.112
		(0.007)	(0.058)	(0.020)	(0.032)
Cross-elasticity, different segment		0.015	0.002	0.082	0.057
		(0.004)	(0.001)	(0.024)	(0.017)
Model selection criteria					
AIC		0	0	0	1000
BIC		0	0	0	1000

The table reports the empirical means and standard deviations (in parentheses) of the parameters, the implied price elasticities for $T = 1$, and two information criteria, GMM-BIC and GMM-AIC. The estimates are based on 1,000 random samples of 50 markets and 25 products. The true model is the RCNL model of setup 2 with the following design: $\sigma_{x^1} = 1.0$ and $\rho = 0.3$; $\varsigma_{xd} = 0.9$; $\gamma = 1$.

for belonging to group 0 or 1 at either $\gamma = 0$ or $\gamma = 1$: in the first case both groups are equally crowded, while in the second case, group 0 is more crowded than group 1. Finally, we set either $\varsigma_{xd} = 0$ or $\varsigma_{xd} = 0.9$: in the first case, the group dummy d_{jt} is not informative about the continuous characteristic x^1_{jt}, while in the second case, d_{jt} is very informative about x^1_{jt}. For each design and each of the 1,000 generated data sets, we estimate the correctly specified RCNL model and the misspecified models (RC with $\rho = 0$, NL with $\sigma_{x^1} = 0$ and logit with $\rho = \sigma_{x^1} = 0$). These experiments enable us to see to what extent ρ takes over the role of σ_{x^1} (and vice versa) and what this implies for the price elasticity estimates under a broad variety of designs.

Table 2 shows detailed results for one of the eight designs: $\sigma_{x^1} = 0.5$ and $\rho = 0.5$, $\gamma = 1$, and $\varsigma_{xd} = 0.9$. We begin with this design because it best mimics our car data set. First, there is relatively strong consumer heterogeneity for the group dummy (high ρ); second, more products belong to group 0 than to group 1 (84% versus 16% of the products); and third, the group dummy is quite informative for the continuous characteristic x^1_{jt} (probit regressions of the group dummy on the continuous characteristic imply 88.4% correct classifications, where group 0 tends to contain the low-price goods and group 1 the high-priced goods).

We first look at the parameter estimates of the correctly specified RCNL model. The means are very close to the true parameters, the standard errors are small, and the distribution is approximately normal (not shown). This again confirms that our estimation procedure works well in practice, as well as in the RCNL model where we used the modified contraction mapping.

The parameter estimates for the logit, NL, and RC models give interesting results on the effects of estimating misspecified models. The NL model, which imposes $\sigma_{x^1} = 0$, leads to an upward bias of the mean valuation of x^1_{jt} and of the nesting parameter ρ: $\beta_{x^1} = -0.31 > -3$ and $\rho = 0.88 > 0.3$. The RC model, which imposes $\rho = 0$, also results in parameter biases: $\beta_{x^1} = -4.06 < -3$ and $\sigma_{x^1} = 1.30 > 1.0$. Hence, the nesting parameter ρ and the standard deviation σ_{x^1} partly take the role of the other omitted parameter.

What do these estimates imply for the substitution patterns? The own-price elasticities tend to be underestimated (in absolute value) in the logit and NL models; in the RC model, they are relatively close to the estimates of the correctly specified RCNL model. The cross-price elasticities between products of the same group are underestimated in all misspecified models (logit, NL, and RC). This confirms the importance of accounting for all sources of consumer heterogeneity (regarding the continuous variable and group dummy). Finally, the cross-price elasticities between products of different groups are underestimated in the NL model and overestimated in the RC model.[14]

These findings refer to one of the eight designs (the design that is closest to our empirical data set below). However, as we show in table 3, our main conclusions remain across all dimensions of the designs. The RC model estimates the own-price elasticities quite well (i.e., close to the ones from the correctly specified RCNL model), whereas the logit and NL models tend to underestimate the own-price elasticities (but the NL model less so when the nesting parameter is strong). The logit, RC, and NL models underestimate substitution within groups, since they do not capture all sources of consumer heterogeneity. The RC model overestimates substitution between different groups, while the NL model underestimates it. Despite these biases, the NL and

[14] Note that the AIC and BIC selection criteria now correctly detect that the data-generating process is the RCNL model in all of the 1,000 generated data sets. Hence, they always lead to the conclusion that one should account for both sources of heterogeneity (on the continuous characteristic and on the group dummy). This differs from setup 1, where the selection criteria often could not distinguish between the NL and RC models (because the group dummy was the only source of heterogeneity).

TABLE 3.—MONTE CARLO RESULTS SETUP 2: OWN- AND CROSS-ELASTICITY UNDER DIFFERENT DESIGNS OF RCNL MODEL

	Logit	NL	RC	RCNL	Logit	NL	RC	RCNL
	$\sigma_{x^1} = 1.0$ and $\rho = 0.3$				$\sigma_{x^1} = 0.5$ and $\rho = 0.5$			
	$\varsigma_{xd} = 0, \gamma = 0$							
Own elasticity	−2.355	−2.598	−5.412	−5.410	−5.761	−5.838	−7.694	−7.797
Cross-elasticity, same segment	0.020	0.064	0.086	0.123	0.041	0.113	0.071	0.144
Cross-elasticity, different segment	0.021	0.005	0.090	0.063	0.043	0.015	0.073	0.035
	$\varsigma_{xd} = 0.9, \gamma = 0$							
Own elasticity	−2.025	−2.383	−5.411	−5.411	−5.685	−5.745	−7.691	−7.721
Cross-elasticity, same segment	0.017	0.096	0.088	0.145	0.043	0.156	0.067	0.200
Cross-elasticity, different segment	0.018	0.005	0.068	0.049	0.045	0.020	0.075	0.037
	$\varsigma_{xd} = 0, \gamma = 1$							
Own elasticity	−2.396	−2.729	−5.380	−5.408	−5.747	−5.818	−7.645	−7.780
Cross-elasticity, same segment	0.025	0.048	0.099	0.114	0.050	0.080	0.082	0.113
Cross-elasticity, different segment	0.016	0.004	0.065	0.044	0.033	0.014	0.056	0.027
	$\varsigma_{xd} = 0.9, \gamma = 1$							
Own elasticity	−2.545	−3.347	−5.205	−5.344	−5.887	−5.879	−7.699	−7.573
Cross-elasticity, same segment	0.028	0.093	0.083	0.112	0.054	0.106	0.083	0.141
Cross-elasticity, different segment	0.015	0.002	0.082	0.057	0.029	0.014	0.060	0.025

The table reports the empirical means and standard deviations (in parentheses) of the price elasticities for $T = 1$. The estimates are based on 1,000 random samples of 50 markets and 25 products. The true model is the RCNL model of setup 2, with eight different designs according to three criteria: (a) $\sigma_{x^1} = 1.0$ and $\rho = 0.3$, or $\sigma_{x^1} = 0.5$ and $\rho = 0.5$; (b) $\varsigma_{xd} = 0$ or $\varsigma_{xd} = 0.9$; or (c) $\gamma = 0$ or $\gamma = 1$.

RC models find stronger substitution within than between groups, in contrast with the logit where substitution is entirely symmetric.[15] These findings will be confirmed in the next sections, where we provide an empirical analysis of the car market.

More generally, these findings stress the importance of accounting for unobserved consumer heterogeneity regarding group dummy variables. But this does not mean that researchers should necessarily estimate an RCNL model when they expect unobserved heterogeneity on group dummy variables. They may also specify an RC model with an additional random coefficient for the group dummy variable. To illustrate this, online appendix B provides results from a Monte Carlo experiment with one random coefficient on a continuous variable and one on a discrete variable. We find that omitting the random coefficient on the discrete variable results in very similar biases as the earlier findings on omitting the nesting parameter from the RCNL model.[16]

[15] The cross-price elasticities in the logit model do not appear to be entirely symmetric in the two lower panels (designs 5–8, with $\gamma = 1$). This is because one of the two groups is less crowded than the other. The cross-elasticities are symmetric if one considers them at a more disaggregate level (e.g., average separately for group 0 and 1).

[16] As a further check, we also compared the Monte Carlo results from the RCNL (last column of table 2) with an RC model with two random coefficients—one on the continuous variable and one on the discrete variable. This model is "misspecified" since it has different distributional assumptions about the discrete variable than the RCNL model. But the implied own-price and cross-price elasticities are very close (respectively −5.727, 0.114, 0.064, compared with −5.344, 0.112 and 0.054 for the RCNL model in table 2). This generalizes our findings in the simpler setup of table 1.

III. Empirical Analysis

A. *Data set for the European Car Market*

We make use of a unique data set on the automobile market maintained by JATO. The data are at the level of the car model (e.g., VW Golf) and include essentially all passenger cars sold during nine years (1998–2006) in nine West European countries. This covers around 90% of the sales in the European Union. The countries are Belgium, France, Great Britain, Germany, Greece, Italy, Portugal, Spain, and the Netherlands. For each model, country, and year, we have information on sales, defined as total new registrations. For models introduced or eliminated within a given year, we know the number of months with positive sales in the given year. We exclude the models with extremely small market shares (e.g., Bentley Arnage or Kia Clarus). This results in a data set of 18,643 model/country/year observations or on average about 230 models per country/year.

We combine the sales data with information on the list prices and various characteristics referring to the base model: vehicle size (curb weight, width, and height), engine attributes (horsepower and displacement), and fuel consumption (liter/100 km or euros /100 km). We start from JATO's classification to assign each model to one of seven possible marketing segments: subcompact, compact, intermediate, standard, luxury, SUV, and sports. Furthermore, we assign the models to their brands' perceived country of origin. For example, the Volkswagen Golf is perceived of German origin even if produced in Spain. We construct a dummy variable for whether a model is of foreign or domestic origin in each country. Our two-level nested logit model will use the marketing segments and foreign origin dummy to define the groups (e.g., subcompact) and subgroups (e.g., domestic

TABLE 4.—SUMMARY STATISTICS

	All countries		France		Germany	
	Mean	SD	Mean	SD	Mean	SD
Sales (units)	5,785	14,694	8,440	19,931	11,432	21,074
Price/income	1.19	0.94	0.90	0.53	0.95	0.63
Horsepower (in kW)	88.8	40.9	87.7	37.4	92.8	44.6
Fuel efficiency (euros /100 km)	8.4	2.1	8.5	2.3	8.8	2.6
Width (cm)	173.0	8.5	173.1	8.5	173.4	8.6
Height (cm)	148.3	13.8	149.2	14.2	148.2	14.1
Foreign (0–1)	0.92	0.28	0.86	0.35	0.71	0.45
Months present (1–12)	9.89	2.55	9.70	2.65	9.77	2.56

Means and standard deviations of the main variables. The total number of observations (models/markets) is 18,643, where markets refer to the nine countries and nine years.

TABLE 5.—SUMMARY STATISTICS BY SEGMENT

Segment	Subcompact	Compact	Intermediate	Standard	Luxury	SUV	Sport
Mean of the characteristics							
Sales (units)	11,155	7,450	5,009	4,632	2,889	2,205	1,517
Price/income	0.55	0.81	1.04	1.39	2.13	1.61	1.85
Horsepower (in kW)	48.7	70.1	84.6	99.6	134.0	113.7	126.6
Fuel efficiency (euros /100 km)	6.4	7.2	8.0	8.7	10.4	11.2	9.6
Width (cm)	162.5	171.4	175.3	175.1	182.3	179.4	175.1
Height (cm)	149.1	144.2	144.9	142.6	145.3	175.9	133.6
Foreign (0–1)	0.92	0.92	0.93	0.91	0.89	0.96	0.86
Months present (1–12)	9.72	9.87	9.88	9.77	9.94	10.11	10.03
Number of observations	3,788	4,095	2,656	1,711	1,764	2,521	2,108
Correct classifications into different marketing segments (in percent)							
Subcompact	–	93.7	99.4	99.9	100.0	95.5	97.6
Compact		–	76.6	91.1	97.7	99.7	92.8
Intermediate			–	77.9	91.4	99.7	91.0
Standard				–	90.0	99.9	84.4
Luxury					–	99.7	88.9
SUV						–	99.9
Sports							–

The top panel reports means of the main variables per segment. The bottom panel reports the percentage of correctly classified car models, based on binary probit of a segment dummy per pair on four continuous characteristics (horsepower, fuel efficiency, width, and height).

subcompact, foreign subcompact). Table 4 provides summary statistics for sales, price, and the product characteristics used in our empirical demand model. We show the summary statistics for all countries and for France and Germany separately (since we will focus on these countries when we present our counterfactuals).

Since our empirical analysis will focus on comparing the nested logit and random coefficients logit models, it is informative to provide background on how the continuous characteristics relate to the marketing segments. Table 5 (top panel) shows summary statistics for our four characteristics by marketing segment. Cars belonging to the same marketing segment tend to have similar horsepower, fuel consumption, width, and height. Horsepower and fuel consumption show a higher dispersion within a segment than width and height, but their segment averages also vary more widely. For example, average horsepower varies from 48.7 kW in the subcompact to 134 kW in the luxury segment, whereas average width varies from 162.5 cm in the subcompact to 182.3 cm in the luxury segment. Table 5 (bottom panel) summarizes how well the four characteristics predict which segment each model belongs to. For each segment pair (e.g., subcompact-compact), we estimate a probit explaining segment assignment as a function of the four characteristics and ask how often the probit correctly classifies the different car models. The table shows that the continuous variables predict the SUV extremely well, with over 95% correct classifications with respect to any other segment. Classification is also quite accurate for most other segments; for example, for the luxury segment, there are over 89% correct classifications with respect to any other segment. The lowest number of correct classifications occurs for a few neighboring segments (on the diagonal), for example, 76.6% correct classifications between compact and intermediate, 77.9% between intermediate and standard. But even in these instances, the characteristics predict the segments quite well.

In sum, this preliminary evidence indicates that a limited number of characteristics (horsepower, fuel consumption, width, and height) have quite good, but not perfect, predictive power for the classification in marketing segments. We will bear this in mind when comparing the NL and RC models.

B. *Specification*

To estimate the logit, NL, RC, and RCNL demand models, we make three modifications to the framework discussed in section II. First, we treat price separately since it is an endogenous characteristic and since we allow its random coefficient to follow the empirical distribution of income. We adopt the following variant of the above utility specification (1):

$$u_{ijt} = x_{jt}\beta_i - \alpha_i p_{jt} + \xi_{jt} + \bar{\varepsilon}_{ijt}.$$

The vector of observed product characteristics, x_{jt}, includes horsepower, fuel efficiency, width, height, and a dummy variable for the product's country of origin (domestic or foreign). The corresponding random coefficients are specified as before: $\beta_{ik} = \beta_k + \sigma_k \nu_{ik}$ for characteristic k. Price p_{jt} enters slightly differently: its random coefficient is specified as $\alpha_i = \alpha/y_i$, where y_i is the income of individual i. In the RC and RCNL model, we treat y_i as a random variable with a known distribution equal to the empirical distribution of income. In the NL model, we treat y_i as nonrandom and set it equal to mean income in market t, $y_i = \bar{y}_t$. In sum, for the nonprice characteristics, we estimate both the mean valuations β_k and the standard deviations σ_k; for price, we estimate only α so that heterogeneity in willingness to pay follows the empirical distribution of income.[17]

Second, the product-specific taste parameter $\bar{\varepsilon}_{ijt}$ follows the distributional assumptions of the two-level nested logit model instead of the one-level nested logit of section II. The upper level consists of the seven different market segments (subcompact, compact, standard, intermediate, luxury, SUV, and sports) and one separate segment for the outside good. The lower level divides every segment in two subsegments according to the models' country of origin (domestic or foreign). In four countries, there are only foreign cars, so the subsegments of domestic cars are empty (Belgium, Greece, Portugal, and the Netherlands). There are now two nesting parameters, $\rho = (\rho_1, \rho_2)$. The nesting parameter ρ_1 measures the correlation of preferences across cars of the same subsegment, and ρ_2 measures correlation of preferences across subsegments of the same segment. For the model to be consistent with random utility maximization, $0 \le \rho_2 \le \rho_1 < 1$. If $\rho_1 = \rho_2$, the model reduces to a one-level nested logit where the segments are the nests; if $\rho_1 > \rho_2 = 0$, the model reduces to a one-level nested logit where the subsegments are the nests. If $\rho_1 = \rho_2 = 0$, the model reduces to a simple logit. Assuming that consumers choose the product that maximizes utility, we obtain a two-level nested logit version of the aggregate market shares, equation (6).

Third, we exploit the panel features of our data set to specify the error term as $\xi_{jt} = \xi_j + \xi_t + \Delta\xi_{jt}$. The ξ_j are product fixed effects, capturing time-invariant unobserved characteristics for a car model j. The ξ_t are market fixed effects, modeled as country-specific fixed effects interacted with a time trend and squared time trend. These capture general country-specific demand shocks relative to the outside good.[18] Finally, $\Delta\xi_{jt}$ is the residual, capturing remaining unobserved product characteristics, varying across products and markets. Since our data are at the annual level, we also include a set of dummy variables for the number of months each model was available in a country within a given year (for models introduced or dropped within a year).

C. *Identification and Estimation*

To estimate the demand parameters $\theta = (\beta, \alpha, \sigma, \rho)$, we follow Berry (1994), BLP, and the subsequent literature. We solve the system $s_t = s_t(\delta_t, \theta)$ for δ_t in each market t to obtain a solution for the error term $\Delta\xi_{jt}$ for each product $j = 1, \ldots, J$ in market t:

$$\delta_{jt}(s_t, \alpha, \sigma, \rho) = x_{jt}\beta + \xi_j + \xi_t + \Delta\xi_{jt}. \tag{9}$$

In the (two-level) NL model, the left-hand side has an analytic solution,

$$\delta_{jt}(s_t, \alpha, \sigma, \rho) = \ln s_{jt}/s_{0t} - \rho_1 \ln s_{j|hgt} - \rho_2 \ln s_{h|gt} + \alpha p_{jt}/\bar{y}, \tag{10}$$

so that a linear estimator can be used. In the RC and RCNL model, $\delta_{jt}(s_t, \alpha, \sigma, \rho)$ should be computed numerically by solving the system $s_t = s_t(\delta_t, \theta)$ for δ_t, which makes estimation considerably more complex.

For all models, we can proceed with GMM by interacting the error term with a vector of instrumental variables z_{jt} that is uncorrelated with the error term. Since there are $2K + 3$ parameters (K mean valuations β_k, K standard deviations σ_k, the price parameter α, and the two nesting parameters ρ_1 and ρ_2), we need at least $2K + 3$ instruments in z_{jt}. Price p_{jt} does not qualify as an instrument to identify the price effect, since it is likely to be correlated with $\Delta\xi_{jt}$. For example, a positive demand shock for product j in market t will not only increase the demand for the product but may also induce the firm to raise its price. Failure to account for this endogeneity issue will lead to an estimated price coefficient (α) that is downward biased.[19] Our identification assumption is that the observed product characteristics x_{jt} are uncorrelated with the unobserved product characteristics $\Delta\xi_{jt}$ (which is weaker than the often adopted assumption that x_{jt} is uncorrelated with ξ_{jt}). As discussed in BLP, one may use alternative functions of these characteristics as instruments to estimate the $2K + 3$ parameters. More specifically, following previous practice, our vector of instrumental variables z_{jt} includes: the vector of product characteristics x_{jt}, the sum of the characteristics of

[17] This utility specification approximates BLP's Cobb-Douglas specification $\alpha \ln(y_i - p_j)$ when the price is small relative to (capitalized) income. It is particularly convenient when studying countries with different exchange rates, since local price is simply expressed relative to local income (see Goldberg & Verboven, 2001).

[18] We set the potential number of consumers L_t as the number of households in the market. Alternative assumptions on L_t are absorbed in the market fixed effects and do not have an important impact on the results, as we discuss below.

[19] In the linear NL model, the within-subgroup and within-group market shares $\ln s_{j|hgt}$ and $\ln s_{h|gt}$ evidently do not qualify as instruments to identify the nesting parameters ρ_1 and ρ_2 (just like functions of market shares would not qualify as instruments for the distributional parameters σ in the RC model).

TABLE 6.—PARAMETER ESTIMATES FOR ALTERNATIVE DEMAND MODELS

	Logit		Nested Logit		RC Logit		RC Nested Logit	
	Parameter	SE	Parameter	SE	Parameter	SE	Parameter	SE
Mean valuations for the characteristics in x_{jt} (β)								
Price/income	−1.76	0.17	−1.00	0.03	−5.52	0.66	−2.75	0.18
Horsepower (kW/100)	2.30	0.24	1.34	0.08	−3.67	1.86	0.57	0.77
Fuel (euros /10,000 km)	−11.48	1.43	−6.13	0.52	−20.77	3.06	−4.68	0.73
Width (cm/100)	2.51	0.55	−0.10	0.29	3.64	0.83	1.26	0.50
Height (cm/100)	3.46	0.35	1.17	0.19	0.27	1.32	2.12	0.46
Foreign (0/1)	−1.21	0.03	−0.47	0.04	−3.66	0.89	−0.57	0.14
Standard deviations of valuations for the characteristics in x_{jt} (σ)								
Horsepower (kW/100)	NA		NA		4.67	0.83	0.92	0.41
Fuel (euros /10,000 km)	NA		NA		1.15	1.69	1.66	0.57
Width (cm/100)	NA		NA		1.93	0.71	0.10	1.74
Height (cm/100)	NA		NA		4.83	0.55	0.15	1.11
Foreign (0/1)	NA		NA		5.46	1.05	0.22	0.84
Constant	NA		NA		1.18	0.43	0.21	3.00
Nesting parameters (ρ_1 and ρ_2)								
Subsegment ρ_1	NA		0.65	0.03	NA		0.57	0.03
Segment ρ_2	NA		0.48	0.03	NA		0.47	0.07
Model fixed effects	Yes		Yes		Yes		Yes	
Market fixed effects	Yes		Yes		Yes		Yes	
Income distribution	No		No		Yes		Yes	
Random coefficients	No		No		Yes		Yes	
# inelastic demands	3,514 (19%)		556 (3%)		0		0	
χ^2 test $\rho_1 = \rho_2$	NA		83.04		NA		2.76	
Prob. χ^2			(0.00)				(0.10)	

Parameter estimates and standard errors for the different demand models. The logit and NL models assume equal income ($-\alpha/\bar{y}_t$), and the RC and RCNL models allow for heterogeneous income ($-\alpha/y_i$). The total number of observations (models/markets) is 18,643, where markets refer to the nine countries and nine years.

other products of competing firms, and the sum of the characteristics of other products of the same firm. For the NL and RCNL models, we also include these sums over products belonging to the same subsegment and segment, following Verboven (1996).[20]

The GMM objective function includes a weighting matrix to account for heteroskedasticity (obtained from the residuals using a two-step procedure). To minimize the GMM objective function with respect to the parameters $\theta = (\beta, \alpha, \sigma, \rho)$, we first concentrate out the linear parameters β (which includes a set of dummy variables for the market fixed effects ξ_t). We do not directly estimate the more than 200 car model fixed effects ξ_j; instead we use a within transformation of the data (Baltagi, 1995). Standard errors are computed using the standard GMM formulas for asymptotic standard errors.

A few recent papers have studied several numerical difficulties with estimating the RC model (which also apply to the RCNL model): global convergence problems and the role of starting values and different optimization algorithms (Knittel & Metaxoglou, 2012), problems with numerically solving δ_t using BLP's contracting mapping (Dubé et al., 2012), and problems with approximating the integral over the logit probabilities using simulation (Judd & Skrainka, 2011).

We draw lessons from this recent literature and proceed as follows. First, to approximate the high-dimensional integral, equation (6), we make use of a large number of Halton draws over the density $N(0, 1)$. This provides a more effective coverage of the density domain than pseudo-random draws. In particular, we take a large number of 500 Halton draws for each of the 81 markets (country/years).[21] Second, to ensure the GMM objective function is smooth, we use a tight tolerance level of $1e^{-12}$ to invert the shares using our modification of BLP's contraction mapping, equation (7). This tolerance level is considerably stricter than typically used in the literature. Third, we program analytic derivatives of the gradient of the objective function. While this is particularly tedious for the RCNL model, it greatly improves accuracy and computation time. Finally, even if the GMM objective function is smooth, it may not be globally convex. To minimize the function with respect to the nonlinear parameters (α, σ, ρ), we use different starting values, using a stringent convergence criterion of $1e^{-6}$ and carefully examining the gradient, the solution path, and the Hessian eigenvalues. We use a BFGS algorithm, an efficient procedure that uses information at different points to obtain a sense of the curvature of the objective function. We usually obtain the same optimum, except for very high or low starting values, but in these cases, the value of the objective function at convergence is always higher.[22]

D. Parameter Estimates

Table 6 shows the parameter estimates for the four demand models. The logit model imposes $\sigma = \rho = 0$ and $y_i = \bar{y}_t$.

[20] Weak instruments tests show that the instruments are jointly significant.

[21] Halton draws can be very effective compared to pseudo-random draws. For example, Bhat (2001) and Train (2000) report that the simulation variance in the estimated parameters is lower with 100 Halton draws than with 1,000 pseudo-random draws.

[22] The log condition number of the Hessian matrix is, at worst, 1.9, which means that only two (of a total of sixteen) decimal places of accuracy are being lost in the calculation of the Hessian, thus suggesting accurate results.

The NL model assumes $\sigma = 0$ and $y_i = \bar{y}_t$ and estimates ρ. The RC model assumes $\rho = 0$, estimates σ, and allows y_i to follow the empirical distribution of income. Finally, the RCNL estimates both ρ and σ, and allows y_i to follow the empirical distribution of income.

In the simple logit model both the price parameter (α) and the mean valuation parameters (β) have the expected signs and are all significantly different from 0. However, as is well known, the model is very restrictive since it imposes symmetric cross-price elasticities. Furthermore, demand is inelastic for almost 20% of the car models across countries and years. This is inconsistent with oligopolistic profit-maximizing behavior unless marginal costs would be negative.

In the NL model, the upper nest level consists of the seven marketing segments, and the lower nest level consists of the segments and origin (domestic or foreign). The price parameter (α) and the mean valuation parameters (β) again have the expected sign and are significantly different from 0, with the exception of the parameter for width, which is now insignificant. The nesting parameters are estimated very precisely, $\rho_1 = 0.65$ and $\rho_2 = 0.48$. Their magnitudes are consistent with the requirements of random utility maximization ($0 \leq \rho_2 \leq \rho_1 < 1$) and imply that consumer preferences show the strongest correlation across cars from both the same marketing segment and origin (domestic or foreign) and show weaker but still important correlation across cars from the same segment but a different origin. This is consistent with earlier work for a more limited set of countries (Goldberg & Verboven, 2001; and Brenkers & Verboven, 2006).[23] As documented below, this implies more plausible cross-price elasticities than the simple logit model. Furthermore, the implied own-price elasticities are higher than in the simple logit: demand is now inelastic for only 3% of the car models. This may seem surprising at first, since the price coefficient α is closer to 0 than in the simple logit model. However, the elasticities depend not only on α but also on the nesting parameters ρ_1 and ρ_2.

In the RC model, we estimate the price parameter (α) and the means (β) and standard deviations (σ) for the valuations of the other characteristics (including the constant). The price parameter (α) is again significantly estimated with the expected sign (negative effect). Consumers have a negative and significant mean valuation for fuel consumption, and heterogeneity is limited so that almost all consumers dislike fuel-inefficient cars. Consumers have a positive and significant mean valuation for width, and the standard deviation implies that about 10% of consumers dislike large cars. Consumers have a negative mean valuation for cars of foreign origin. The standard deviation is relatively large, so that 25% of consumers actually prefer foreign cars. The mean valuation for height is insignificantly different from 0, and the mean valuation for horsepower is unexpectedly negative. However, for both characteristics, we find substantial and significant heterogeneity: about 50% of consumers have a positive valuation for height, and about 30% have a positive valuation for horsepower. Finally, we estimate a significant standard deviation for the constant, indicating significant heterogeneity in the valuation of new cars relative to the outside good. Overall, the random coefficients show evidence of significant consumer heterogeneity in several dimensions, in particular height, horsepower, and foreign origin. Yet it is striking that the random coefficients are estimated much less precisely than the two nesting parameters in the NL model.

In the RCNL model, we combine the previous two models, so we include both the nesting parameters and the random coefficients. Both the price parameter (α) and the mean valuation parameters (β) have the expected signs and are estimated significantly with the exception of the horsepower parameter, which is insignificant. The most interesting findings relate to the estimated nesting parameters (ρ) and random coefficients (σ) in comparison with the NL and RC models.

First, compared with the NL model, the nesting parameters remain highly significant, but their magnitude becomes smaller. This is consistent with the results from our Monte Carlo study, where we found an overestimate of the nesting parameters if the random coefficients are important and the groups are correlated with the characteristics for the omitted random coefficients. Furthermore, we can no longer reject the hypothesis that $\rho_1 = \rho_2$ (P-value 0.0967) and the random coefficient for foreign origin is insignificant. So the model reduces to a one-level nested logit with no need to divide the seven segments into domestic and foreign subgroups, and it seems at first that there is no longer consumer heterogeneity for foreign origin. However, the subsegment parameter ρ_1 captures similar effects as the random coefficient for foreign origin, suggesting it is not sensible to include both. Indeed, in a one-level nested logit where we constrain $\rho_1 = \rho_2$ (so that the subgroups are no longer relevant), the random coefficient for foreign origin becomes significant again (as in the RC model). We show these results in online appendix B.[24]

Second, compared with the RC model, the random coefficients for horsepower and fuel efficiency remain significant, but this is no longer the case for width, height, and the constant. Intuitively, the nesting parameter for the segments captures a lot of the heterogeneity relating to the car dimensions and the outside good, but not much of the heterogeneity relating to horsepower and fuel efficiency.

Since the logit, NL, and RC are all restricted versions of the RCNL model, we can compare their statistical performance

[23] We also estimated a two-level NL model with the reverse nesting structure, where origin defines the upper level and origin or segment of the lower level of the nests. This led to estimates of ρ_1 and ρ_2 inconsistent with random utility maximization, in line with the results of other studies on the car market.

[24] In this case, the one-level nested logit with a random coefficient for foreign origin seems preferable to a two-level nested logit model, since it does not impose the consumer heterogeneity to enter in a hierarchical way. Nevertheless, we base our subsequent discussion on the two-level nested logit. The implied price elasticities and competition policy counterfactuals are very similar in the one-level nested logit model (not shown).

TABLE 7.—LIKELIHOOD RATIO TESTS FOR ALTERNATIVE DEMAND MODELS

	Logit	Nested Logit	RC Logit
Logit	–		
Nested logit	584.08	–	
	(0.0000)		
RC logit	34.08	NA	–
	(0.0000)		
RC nested logit	534.10	30.61	423.84
	(0.0000)	(0.0002)	(0.0000)

χ^2 statistics and P-values (in parentheses) of likelihood ratio tests for different model pairs.

using likelihood ratio tests adapted to the GMM context.[25] Table 7 reports LR values and asymptotic P-values for all pairs of models, except the NL and RC, which are not nested in each other. Each restricted model is rejected against the more general models. The logit is clearly rejected against any other model. More interesting, both the NL and RC models are rejected against the more general RCNL model. In fact, the NL appears to provide a better fit than the RC logit relative to the RCNL, since the χ^2 statistic is lower for the NL than the RC model (30.61 versus 423.84). We already observed that the individual random coefficients in the RC model are much less precisely estimated than the two nesting parameters in the NL model. The likelihood ratio tests thus indicate that the random coefficients of the RC model are also jointly less significant than the nesting parameters of the NL model.

Summary. We can summarize our empirical results in four points:

a. It is important to include the nesting parameter relating to the seven marketing segments since it remains highly significant after including the random coefficients.
b. It does not seem appropriate to include an additional subnesting parameter relating to the origin within each segment, since the random coefficient for origin captures this well.
c. It is relevant to include random coefficients for horsepower and fuel efficiency, but not those for the dimensions width and height since these are captured well by the marketing segments.
d. It is striking that the nesting parameters (reflecting heterogeneity regarding segments and subsegments) are estimated much more precisely than the random coefficients (reflecting heterogeneity regarding continuous characteristics).

While these findings apply to our data set of the European car market, they can be useful as a guide for interpretations also in other applications.

[25] Following Hayashi (2000), we define the likelihood ratio statistic (*LR*) as the difference between the value of the objective function of the restricted model (reestimated using the second-stage weighting matrix of the unrestricted model) and the value of the objective function of the unrestricted model. Under the null hypothesis, the statistic is asymptotically χ^2 distributed with degrees of freedom equal to the number of restrictions.

We have done various sensitivity analyses to assess the robustness of these conclusions. We added a random coefficient for the continuous characteristic weight, or dropped the random coefficient for height. In both cases, the magnitude and significance of the nesting parameter for the segments are hardly affected.[26] We also estimated a model with an additional upper nest to distinguish the three lower class segments (subcompact, compact, and standard) from the other four segments. We did not find this to be an additional source of segmentation. In sum, this suggests that the segments capture a separate, unobserved source of market segmentation. In principle, this may be captured without imposing a nested logit structure, through random coefficients on the segment dummies with another distribution (e.g., normal as for the continuous variables). This would, however, substantially increase the computational burden.

E. *Substitution Patterns*

We have already commented on the number of inelastic own-price elasticities implied by our estimates. We now provide a more systematic discussion on the substitution patterns. We consider own-price and cross-price elasticities at the level of the individual products and at the level of the entire segments.

Product-level price elasticities. First, consider the product-level own- and cross-price elasticities. We average these by segment and distinguish between cross-price elasticities with respect to other products in the same subsegment, in a different subsegment within the same segment, and in a different segment. Table 8 shows these average product-level elasticities for Germany in 2006 (the largest country in the most recent year of our data set). In the logit and NL models, the own-price elasticities tend to increase more or less proportionally with price as one moves to higher segments, resulting in an average own-price elasticity that is almost four times higher in the luxury than in the subcompact segment. The near proportional relationship follows from the functional form assumption: price enters utility linearly with a homogeneous valuation across consumers ($-\alpha/\overline{y}_t$). In contrast, in the RC and RCNL models, the price elasticities increase much less than proportionally, by a factor of 2.2 and 2.3 in the respective models. This follows from the less restrictive functional form: price still enters utility linearly, but consumer valuations are heterogeneous ($-\alpha/\overline{y}_{it}$). Hence, price-insensitive consumers are more likely to purchase high-priced cars.

[26] Adding random coefficients raises the computational burden of approximating the market share integral and makes it more difficult to identify the standard deviations of the individual random coefficients. To reduce the number of random coefficients, we also conducted a principal components analysis before estimating the model. We find that there are two main principal components: the first is closely related to the performance and size variables; the second is closely related to the sports aspect (acceleration and height). We estimate a strongly significant random coefficient for the first principal component and a less significant one for the second component.

TABLE 8.—PRODUCT-LEVEL PRICE ELASTICITIES IN GERMANY FOR ALTERNATIVE DEMAND MODELS

	Own-	Cross-Price Elasticity		
Segment	Price Elasticities	Same Subsegment	Same Segment	Different Segment
Logit				
Subcompact	−0.76	<0.01	<0.01	<0.01
Compact	−1.09	<0.01	<0.01	<0.01
Intermediate	−1.49	<0.01	<0.01	<0.01
Standard	−1.94	<0.01	<0.01	<0.01
Luxury	−2.94	<0.01	<0.01	<0.01
SUV	−2.32	<0.01	<0.01	<0.01
Sports	−2.73	<0.01	<0.01	<0.01
Nested logit				
Subcompact	−1.23	0.02	0.01	<0.01
Compact	−1.74	0.03	0.02	<0.01
Intermediate	−2.38	0.05	0.03	<0.01
Standard	−3.04	0.13	0.05	<0.01
Luxury	−4.64	0.17	0.07	<0.01
SUV	−3.73	0.05	0.04	<0.01
Sports	−4.40	0.08	0.03	<0.01
RC logit				
Subcompact	−2.85	0.03	<0.01	<0.01
Compact	−3.66	0.02	<0.01	0.01
Intermediate	−4.38	0.03	<0.01	0.01
Standard	−4.96	0.04	0.01	0.01
Luxury	−6.24	0.06	0.03	0.01
SUV	−5.67	0.04	<0.01	0.01
Sports	−6.13	0.02	<0.01	0.02
RC nested logit				
Subcompact	−2.57	0.03	0.03	<0.01
Compact	−3.33	0.05	0.05	<0.01
Intermediate	−3.90	0.06	0.06	<0.01
Standard	−4.54	0.15	0.09	<0.01
Luxury	−5.75	0.17	0.11	<0.01
SUV	−5.01	0.07	0.06	<0.01
Sports	−5.42	0.10	0.05	<0.01

Product-level own- and cross-price elasticities, based on the parameter estimates in table 6. Elasticities are averages by segment for Germany in 2006. Cross-price elasticities are averaged across products from the same subsegment, from a different subsegment within the same segment, and from different segments.

The cross-price elasticities show even more striking differences across the estimated models. In the logit model, they are extremely small even with respect to cars from the same subsegment or segment (always less than 0.01). In contrast, in the NL and RCNL models, the cross-price elasticities are quite high with respect to products of the same subsegment (about 0.1–0.4), and they are still relevant with respect to products of other subsegments in the same segment (about 0.05). In the RC model, the cross-elasticities with respect to products of the same subsegment are still sizable, mainly because of the magnitude and significance of the foreign ownership random coefficient. But they are negligible with respect to products of other segments within the same segment (usually below 0.01). These findings illustrate the importance of accounting for consumer heterogeneity relating to the marketing segments (as done only in the NL and RCNL models) and the domestic or foreign origin (as done in all models except the simple logit).

Segment-level price elasticities. Now consider the segment-level price elasticities, that is, the effect of a joint 1% price increase of all cars in a given segment on demand in the various segments. Table 9 reports these segment-level own- and cross-price elasticities. We can summarize these results as follows. First, as is well known, both the logit and NL models imply fully symmetric substitution patterns at the segment level (i.e., identical cross-elasticities per row). For example, a price increase of all compact cars by 1% raises the demand in all other segments by 0.02% (more precisely, by 0.017%). In sharp contrast, the RC model implies more intense substitution to neighboring segments. Taking the same example, a price increase of all compact cars by 1% has the highest effect on the demand for subcompact (+0.76%) and intermediate cars (+0.66%) and the lowest effects on the demand for luxury (0.26%) and SUV cars (+0.39%). Finally, the RCNL model implies cross-price elasticities somewhere in between the NL and RC models, though closer to the NL model: the cross-price elasticities to other segments are fairly (but not completely) symmetric, and they are somewhat higher than in the NL model, but not nearly as high as in the RC model.

Although the substitution patterns of the most general RCNL model appear to be better approximated by the NL model than by the RC model, this does not necessarily mean that the NL model should be preferred over the RC model. The main message is that it is important to account for consumer heterogeneity regarding the marketing segments. The NL model is one simple way to capture this, but there may be alternative ways. For example, one may consider adding random coefficients for the segments at an increased computational cost.

Summary. We can summarize the differences in the estimated substitution patterns across models as follows. First, the own-price elasticities at the product level increase roughly proportionally with price in the logit and NL models, but less than proportionally in the RC and RCNL models. This is because the latter two models allow consumer heterogeneity in the price parameter. Second, the product-level cross-price elasticities show that products of the same segment are strong substitutes in the NL and RCNL models, but not in the logit and RC models. Finally, the segment-level cross-price elasticities show quite strong substitution across segments (especially the neighboring ones) in the RC model but only weak (and symmetric) substitution in the logit, NL, and RCNL models.

IV. Implications for Competition Policy Analysis

Section III showed how the different demand models generate quite different substitution patterns. But how relevant are the found differences for applications in industrial organization or related fields? To address this question, we consider two areas of competition policy, market definition and merger simulation, and ask whether the different demand models yield robust conclusions.

Much of competition policy still heavily relies on market definition and an assessment of the firms' market shares within the defined market. It is simple and widely applicable to mergers and horizontal or vertical agreements because it

TABLE 9.—SEGMENT-LEVEL PRICE ELASTICITIES IN GERMANY FOR ALTERNATIVE DEMAND MODELS

Segment	Subcompact	Compact	Intermediate	Standard	Luxury	SUV	Sport
Logit							
Subcompact	−0.77	0.02	0.02	0.02	0.02	0.02	0.02
Compact	0.02	−1.12	0.02	0.02	0.02	0.02	0.02
Intermediate	0.01	0.01	−1.41	0.01	0.01	0.01	0.01
Standard	0.01	0.01	0.01	−1.75	0.01	0.01	0.01
Luxury	0.01	0.01	0.01	0.01	−2.59	0.01	0.01
SUV	0.01	0.01	0.01	0.01	0.01	−2.24	0.01
Sports	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	−2.05
Nested logit							
Subcompact	−0.44	0.01	0.01	0.01	0.01	0.01	0.01
Compact	0.01	−0.64	0.01	0.01	0.01	0.01	0.01
Intermediate	<0.01	<0.01	−0.81	<0.01	<0.01	<0.01	<0.01
Standard	<0.01	<0.01	<0.01	−1.00	<0.01	<0.01	<0.01
Luxury	<0.01	<0.01	<0.01	<0.01	−1.48	<0.01	0.01
SUV	<0.01	<0.01	<0.01	<0.01	<0.01	−1.28	<0.01
Sports	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	−1.17
RC logit							
Subcompact	−1.72	0.67	0.47	0.19	0.07	0.33	0.29
Compact	0.75	−2.77	0.66	0.54	0.26	0.39	0.41
Intermediate	0.29	0.39	−3.47	0.43	0.30	0.45	0.42
Standard	0.12	0.32	0.44	−3.55	0.56	0.43	0.45
Luxury	0.05	0.16	0.32	0.61	−4.05	0.86	0.67
SUV	0.15	0.18	0.37	0.43	0.92	−4.13	0.75
Sports	0.08	0.11	0.20	0.25	0.42	0.49	−4.36
RC nested logit							
Subcompact	−1.08	0.04	0.04	0.04	0.04	0.04	0.04
Compact	0.04	−1.42	0.05	0.06	0.06	0.06	0.05
Intermediate	0.03	0.03	−1.65	0.04	0.04	0.04	0.04
Standard	0.03	0.03	0.04	−1.90	0.05	0.05	0.05
Luxury	0.03	0.04	0.05	0.06	−2.37	0.08	0.07
SUV	0.03	0.04	0.05	0.06	0.08	−2.12	0.07
Sports	0.02	0.02	0.03	0.03	0.04	0.04	−2.03

The segment-level own- and cross-price elasticities (when all products in the same segment raise their price by 1%), based on the parameter estimates in table 6. The elasticities refer to Germany in 2006.

makes few assumptions about oligopoly behavior. However, the choice of candidate-relevant markets can often be quite arbitrary and artificial. Furthermore, because it is not based on a specific model of oligopoly behavior, it cannot make precise predictions about market power effects or incorporate other considerations in an integrated framework. In merger cases, one increasingly resorts to simulation analysis to assess market power effects and incorporate efficiencies or other elements (see Werden & Froeb, 1994; Hausman, Leonard, & Zona, 1994; Nevo, 2000; Peters, 2006). While merger simulation may in principle extend to other types of competition investigations, this is difficult in practice because it requires the specification of an appropriate oligopoly model for the specific competition issue under investigation.

These relative advantages and disadvantages of market definition and merger simulation have been widely discussed. We will instead look at this from a different angle; we ask to which extent both approaches are sensitive to the adopted demand model. If one approach gives more robust conclusions across demand models, this provides a new motivation to prefer it over the other approach.

A. *Market Definition*

Market definition in the European car market is relevant not only for the evaluation of mergers, but also for the implementation of the block exemption regulation for the selective and exclusive distribution system. According to this regulation, automobile manufacturers may impose selective or exclusive distribution to their dealers, provided they have market shares below 30% or 40%. Some niche manufacturers such as Mercedes or BMW may meet these thresholds if markets are defined widely to include all cars, but not if they are defined narrowly. Hence, it is important to know the size of the relevant markets.

According to the small but significant and nontransitory increase in price (SSNIP) test, the relevant market is the smallest group of products for which a hypothetical monopolist could profitably impose a small, nontransitory but significant increase in price (typically 5%–10%). Since the profitability of a price increase depends on the extent of substitution to other goods, the estimated demand model is of central importance. For each of the four estimated demand models, we first compute all products' implied marginal costs assuming multiproduct price-setting firms (following BLP, Nevo, 2000, and others). Given the estimated demand systems and the marginal costs, we then ask whether a 10% price increase by all products in a candidate-relevant market raises profits.

We begin with considering the marketing segments as the candidate-relevant markets. Table 10A shows the SSNIP test results for France and Germany in 2006. The logit model suggests that none of the seven marketing segments can be considered as separate relevant markets. For example, a joint

TABLE 10.—RELEVANT MARKET DEFINITION IN FRANCE AND GERMANY

	Logit		Nested Logit		RC Logit		RC Nested Logit	
	France	Germany	France	Germany	France	Germany	France	Germany
A: Candidate Markets Are Segments								
Subcompact	−0.1	−0.2	5.0	6.7	4.5	4.9	8.8	11.0
Compact	−0.6	−0.5	7.2	8.7	−5.1	−1.4	10.8	12.6
Intermediate	−1.0	−1.0	7.4	8.4	−8.6	−5.3	10.4	10.4
Standard	−1.6	−1.5	13.5	11.1	−7.8	−5.1	16.3	13.3
Luxury	−3.4	−3.2	16.2	15.0	−9.5	−5.9	16.6	15.2
SUV	−2.4	−2.6	16.5	15.7	2.9	−5.9	18.1	16.0
Sports	−1.4	−2.4	10.1	13.9	−11.2	−9.1	12.6	14.2
B: Candidate Markets Are Selected Products' Ten Closest Substitutes								
VW Polo	−0.4	−1.0	2.0	3.4	1.4	2.6	3.3	4.0
VW Golf	−0.4	−1.0	4.5	6.8	−1.6	2.6	6.0	8.3
VW Passat	−0.4	−1.0	4.0	5.1	−1.4	−0.5	4.8	5.5
Audi A4	−0.4	−1.0	10.8	10.1	−3.5	−1.3	12.3	11.7
Audi A6	−0.4	−1.0	11.8	13.5	−3.3	−2.3	10.7	13.1
BMW X3	−0.4	−1.1	5.5	7.7	−4.6	−1.6	2.0	4.3
Mercedes SLK Class	−0.4	−1.1	4.0	5.0	−5.2	−2.3	3.4	1.1

Percentage profit increases implied by a joint 10% price increase of all products in the same segment (panel A) and for selected products' ten closest substitutes (panel B). The results are based on the parameter estimates in table 6, assuming marginal costs implied by multiproduct Bertrand competition. The effects refer to France and Germany in 2006.

10% price increase in the compact segment in France reduces profits by 0.6%. The RC model yields a similar conclusion: only the subcompact segment can be defined as a relevant market in both France and Germany. In sharp contrast, the NL and RCNL models result in higher and positive profit effects, implying that all marketing segments constitute separate relevant markets. A joint 10% price increase in the compact segment in France would raise profits by 7.21% according to the NL model and even by 10.84% according to the RCNL model. This narrow market definition follows, of course, from the high significance of the nesting parameter for the segments in the NL and RCNL models.[27]

Should we conclude that the RC model fails to define the markets narrowly at the segment level, in contrast with the more general RCNL model against which it was rejected? The answer may seem to be yes, since we found that the RC model omits important unobservables relating to the marketing segments that are captured in the more general RCNL model. However, proper caution is warranted. First, the RCNL model is itself restrictive since it imposes largely symmetric substitution across the segments. As an alternative to the RCNL model, one may also include random coefficients for segment dummies within an RC framework (at an increased computational cost, since it requires approximating a higher dimensional market share integral). Second, even an RC model without random coefficients on segment dummies may result in a narrower market definition if we do not restrict attention to segments as candidate-relevant markets, but instead consider the set of nearest substitutes (which are more likely to include products from other segments in an RC model).

To assess this second possibility, we considered candidate-relevant markets according to the products' ten closest substitutes, based on the estimated cross-price elasticities for the four different demand models. Table 10B implements this for seven representative products, the top-selling cars in Germany in each segment (e.g., the VW Golf in the compact segment or the Audi A4 in the standard segment). As may be expected, for the NL and RCNL, these ten closest substitutes almost always come from the same segment. In contrast, for the RC model, they often come from other segments since the model does not explicitly account for segments as a source of differentiation.[28] Interestingly, despite the fact that we now include closer substitutes in the RC model, the conclusions from Table 10B remain similar to those in Table 10A. The relevant market is not wider than each product's ten closest substitutes under the NL and RCNL models, while it is always wider under the logit model. Under the RC model, it is also usually wider, though there are two exceptions: for the VW Polo, the relevant market is as narrow as its ten closest substitutes in both France and Germany; for the VW Golf, the relevant market is also narrow in Germany, though it is wider in France. We also computed the minimum number of closest substitutes to form a separate relevant market for the four different demand models. In general, we found that the fewest number of products has to be included for the NL and RCNL model (about five to ten vehicles), a larger number for the RC model (about ten to fifteen vehicles), and the largest for the simple logit model.

In sum, when we use the closest substitutes as a selection criterion for defining candidate markets, the RC model results in a narrower market definition than when we use the segment

[27] Market definition may be sensitive to the definition of the potential market size. Recall that we specified the potential number of consumers L_t as the number of households. We reestimated the demand models by scaling L_t up or down by a factor of 2 or 4. The SSNIP test conclusions are robust: the logit and RC model still predict a wide market definition, while the NL and RCNL models predict the reverse. To illustrate, we report the SSNIP test results for $L_t/2$ in online Appendix B.

[28] To illustrate, online appendix B reports the ten closest competitors for the VW Golf for each estimated demand model. In the logit model, the closest competitors are simply the top-selling cars; in the RC model, they are cars with similar characteristics, which may come from different segments. In the NL and RCNL models, the ten closest competitors all come from the same segment (with one exception).

TABLE 11.—EFFECTS OF TWO HYPOTHETICAL MERGERS IN FRANCE AND GERMANY

	All	Subcompact	Compact	Intermediate	Standard	Luxury	SUV	Sport
				PSA-Renault Merger in France				
Domestic market shares (%)								
PSA	33.4	35.3	38.8	46.0	–	19.1	–	37.3
Renault	22.7	29.8	20.9	17.8	–	9.5	–	13.5
Predicted domestic price increase (%)								
Logit	0.9	1.6	0.9	0.75	0.0	0.2	0.0	0.5
Nested logit	15.5	31.2	13.5	12.8	0.0	2.1	0.0	7.0
RC logit	20.2	37.1	22.6	24.1	0.6	4.8	0.1	14.0
RC nested logit	8.3	15.9	8.0	8.2	−0.1	1.5	−0.1	4.5
				VW-BMW Merger in Germany				
Domestic market shares (%)								
BMW	10.6	2.1	7.9	–	39.6	25.3	15.2	10.8
VW	30.8	23.1	36.3	53.8	31.3	32.4	12.0	21.4
Predicted domestic price increase (%)								
Logit	0.3	0.3	0.4	0.3	0.6	0.3	0.2	0.2
Nested logit	2.9	0.6	2.8	0.1	10.0	4.3	1.6	1.1
RC logit	2.2	0.6	2.0	1.8	4.9	3.2	1.7	1.5
2RC nested logit	1.9	0.6	1.8	0.5	5.8	3.0	1.1	0.9

Percentage price increases for two hypothetical mergers, PSA-Renault and BMW-VW, in their domestic markets, France and Germany, based on the parameter estimates in table 6 and assuming multiproduct Bertrand competition. The effects refer to France and Germany in 2006. Confidence intervals of 95%, based on a bootstrapping procedure, are shown in online appendix B. For example, the 95% confidence interval for the overall predicted price increase after the PSA-Renault merger is [0.7–1.8]% for the logit, [12.5–18.3]% for the NL, [14.6–27.2]% for the RC, and [5.4–15.7]% for the RCNL model.

criterion, but it is still wider than the market definition in the NL and RCNL models. This does not mean that the RC model necessarily leads to a wider market definition than the NL or RCNL model. It just stresses the importance of incorporating sufficient sources of heterogeneity in the RC model, in particular, random coefficients for the segments. The nesting parameters are one way to achieve this, but within an RC model, it is also possible to make other distributional assumptions.

In practical terms, using the closest substitutes to define markets may become tedious, especially in terms of presenting unambiguous, nonoverlapping market definitions in competition cases. As a simpler alternative, one may define two neighboring segments as the relevant market in the RC model (as suggested by the cross-price elasticities). Our SSNIP test results at the level of neighboring segments (not shown) demonstrate that two neighboring segments still do not form relevant markets in the logit model, but they do form relevant markets in the RC model: a joint 10% price increase would raise profits for compact + intermediate (+1.6%), for example, but not for compact + luxury (–1.2%).

B. *Merger Simulation*

We consider the effects of two hypothetical mergers. The first merger is between the two French manufacturers PSA (Peugeot and Citroën) and Renault, and the second merger is between the two German manufacturers BMW and Volkswagen (Volkswagen, Audi, Seat, and Skoda). As shown in table 11, PSA and Renault are strong in their home market, France, with a combined market share of 56% (mainly due to the mass segments). BMW and Volkswagen are slightly less strong in their home market, Germany, with a combined market share of 41%. But they have a particularly strong presence in the standard segment (71%) and the luxury segment (58%).

We first compute the products' marginal costs assuming multiproduct price-setting firms, as we also did to implement market definition. Given the estimated demand systems and the marginal costs, we then predict the new Nash equilibrium resulting from the changed ownership structure after the merger. Intuitively, a merger will entail high price effects if the merging firms sell close substitutes with respect to each other (low cross-price elasticities) and weak substitutes with respect to outsider firms (low own-price elasticities).

Table 11 shows the predicted price effects of the two mergers in the firms' home markets. We also briefly comment on the effects in the foreign markets and show these results in online appendix B. We show the percentage price increases for both the entire market and each of the seven marketing segments (using price indices, where postmerger market shares are the weights).

For both mergers, the logit model predicts very small domestic price effects, despite the merging firms' strong domestic market presence. In sharp contrast, the NL, RC, and RCNL models give more robust conclusions. The PSA-Renault merger would result in large aggregate price increases in France (between 8.3% and 20.2%). The overall predicted price increases are closeest for the NL and RC models (15.5% and 20.1%). They are somewhat lower for the RCNL model (20.2%), but the confidence intervals still overlap (as shown in online appendix B). The BMW-VW merger entails more modest price increases in its home country, Germany, but the results are again robust across all models except the logit model (between 1.9% and 3.0%). In particular, the predicted price increases are the largest in the standard segment, where the German producers have the strongest presence (between 4.9% and 10.0%). While the NL, RC, and RCNL all give fairly robust conclusions regarding the predicted merger effects, the NL model gives more precise predictions than the RC model (as shown by the smaller confidence intervals in online appendix B). The predictions from the RC model consequently also show more sensitivity under alternative specifications.[29]

The predicted price effects in the foreign markets are much smaller (shown in online appendix B). But there is again a notable difference between the logit model and the other three models (where the predicted effects are between 0.4% and 0.6% for the BMW-VW merger in France and between 0.2% and 0.4% for the PSA-Renault merger in Germany).

In sum, from a practical perspective, these findings show that it is clearly inappropriate to use a simple logit model with its symmetric substitution patterns. The merger predictions from the NL, RC, or RCNL model are broadly consistent, although the confidence intervals are higher in the RC model (so that they are also more sensitive to variations in the specification).

C. *Summary*

We can summarize our findings on market definition and merger simulation as follows. Merger simulation yields fairly clear conclusions across different demand models: the simple logit model is clearly inappropriate, but a generalization to the NL, RC, or RCNL gives fairly robust conclusions (though less precise for the RC model). In contrast, market definition depends more heavily on the adopted demand model. In particular, the RC model suggests a wider market definition than the NL and RCNL models, which directly incorporate the segments as a segmentation source.

V. Conclusion

We started from an aggregate RCNL model to provide a systematic comparison between the simple logit and NL models and the computationally more complex RC model. We first used simulated data to document parameter biases from estimating an NL or RC model. We then use data on the automobile market to estimate the different models and, as an illustration, to assess what they imply for competition policy analysis. Our main findings on the advantages and disadvantages of the NL and RC models can be summarized as follows.

In terms of statistical performance, both the NL and the RC models are rejected against the RCNL model. The NL model appears to be less strongly rejected (much lower χ^2) than the RC model, and the nesting parameters of the NL model (ρ) drop by only a modest amount after including random coefficients on continuous variables (σ) in the RCNL model. Furthermore, the nesting parameters are estimated more precisely than the random coefficients. This shows that the marketing segments capture an important, separate source of unobserved consumer heterogeneity. In principle, this could be captured with random coefficients in the RC model, but this would come at an increased computational burden.

In terms of substitution patterns, the NL and RC models yield quite different results. The own-price elasticities increase nearly proportionally with price in the NL model and less than proportionally in the RC model, because the latter model allows for consumer heterogeneity in the price parameter. Furthermore, products within the same segment are much closer substitutes in the NL model, whereas there is strong substitution to other segments (especially to neighboring ones) in the RC model.

Despite the rather different substitution patterns, the NL and RC models generate quite robust conclusions on the predicted price effects from mergers. In sharp contrast, the conclusions for market definition are not robust: markets are defined more narrowly in the NL and RCNL models than in the RC or logit models. This suggests two implications for competition policy. First, in market definition, it is important to directly account for the segment dummies as direct sources of market segmentation. Second, in merger simulation, the conclusions are more robust across demand models, suggesting that the simple NL model can be sufficient to obtain reliable policy conclusions despite the different substitution patterns.

More generally, one can draw two implications for the choice of demand model in applied work. First, the choice between the tractable NL model and the computationally more complex RC model may depend on the application. In our merger analysis, we considered two domestic mergers. In this case, a particularly relevant aspect of consumer heterogeneity is the cars' domestic or foreign origin, which the NL model captures reasonably well. In other applications, the most relevant aspects of consumer heterogeneity may not be captured well by nesting parameters for groups or subgroups. In these cases, it is appropriate to estimate RC models with random coefficients for the most relevant continuous characteristics.

Second, our findings show that it is important to account for sources of market segmentation that are not captured by the continuously measured characteristics in the RC model. We established this by adding a nested logit structure to BLP's random coefficients model (which is computationally simpler than adding random coefficients for the segment dummies with other distributions). In future research, one may also consider other tractable models from the GEV family to capture additional sources of heterogeneity.

[29] We dropped the random coefficient for height and added one for weight as alternative specifications. This results in lower predicted price effects in the RC model, but the patterns across the different segments remain comparable (see online appendix B).

REFERENCES

Andrews, Donald W. K., "Consistent Moment Selection Procedures for Generalized Method of Moments Estimation," *Econometrica* 67 (1999), 543–564.

Armstrong, Timothy B., "Large Market Asymptotics for Differentiated Product Demand Estimators with Economic Models of Supply," Stanford University technical report (2012).

Baltagi, Badi H., *Econometric Analysis of Panel Data* (New York: John Wiley, 1995).

Berry, Steven T., "Estimating Discrete-Choice Models of Product Differentiation," *RAND Journal of Economics* 25 (1994), 242–262.

Berry, Steven T., and Ariel Pakes, "Comments on Alternative Models of Demand for Automobiles, by Charlotte Wojcik," *Economics Letters* 74 (2001), 43–51.

Berry, Steven T., James Levinsohn, and Ariel Pakes, "Automobile Prices in Market Equilibrium," *Econometrica* 63 (1995), 841–890.

——— "Voluntary Export Restraints on Automobiles: Evaluating a Trade Policy," *American Economic Review* 89 (1999), 400–430.

Bhat, Chandra R., "Quasi-Random Maximum Simulated Likelihood Estimation of the Mixed Multinomial Logit Model," *Transportation Research Part B: Methodological* 35 (2001), 677–693.

Brenkers, Randy, and Frank Verboven, "Liberalizing a Distribution System: The European Car Market," *Journal of the European Economic Association* 4 (2006), 216–251.

Bresnahan, Timothy F., "Departures from Marginal-Cost Pricing in the American Automobile Industry: Estimates for 1977–1978," *Journal of Econometrics* 17 (1981), 201–227.

Bresnahan, Timothy F., Scott Stern, and Manuel Trajtenberg, "Market Segmentation and the Sources of Rents from Innovation: Personal Computers in the Late 1980s," *RAND Journal of Economics* 28 (1997), S17–S44.

Cardell, Scott N., "Variance Components Structures for the Extreme-Value and Logistic Distributions with Application to Models of Heterogeneity," *Econometric Theory* 13 (1997), 185–213.

Chamberlain, Gary, "Asymptotic Efficiency in Estimation with Conditional Moment Restrictions," *Journal of Econometrics* 34 (1987), 305–334.

Davis, Peter, and Pasquale Schiraldi, "The Flexible Coefficient Multinomial Logit (FC-MNL) Model of Demand for Differentiated Products," mineograph (July 2012).

Dubé, Jean-Pierre H., Jeremy T. Fox, and Che-Lin Su, "Improving the Numerical Performance of Static and Dynamic Aggregate Discrete Choice Random Coefficients Demand Estimation," *Econometrica* 80 (2012), 2231–2267.

Feenstra, Robert C., and James A. Levinsohn, "Estimating Markups and Market Conduct with Multidimensional Product Attributes," *Review of Economic Studies* 62 (1995), 19–52.

Fershtman, Chaim, and Neil Gandal, "The Effect of the Arab Boycott on Israel: The Automobile Market," *RAND Journal of Economics* 29 (1998), 193–214.

Goldberg, Pinelopi Koujianou, and Frank Verboven, "The Evolution of Price Dispersion in the European Car Market," *Review of Economic Studies* 68 (2001), 811–848.

Hausman, Jerry A., Gregory Leonard, and J. Douglas Zona, "Competitive Analysis with Differentiated Products," *Annales d'Economie et de Statistique* 34 (1994), 159–180.

Hayashi, Fumio, *Econometrics* (Princeton, NJ: Princeton University Press, 2000).

Judd, Kenneth L., and Ben Skrainka, "High Performance Quadrature Rules: How Numerical Integration Affects a Popular Model of Product Differentiation," CeMMAP working papers CWP03/11 (February 2011).

Knittel, Christopher R., and Konstantinos Metaxoglou, "Estimation of Random Coefficient Demand Models: Two Empiricists' Perspective," this REVIEW (forthcoming).

McFadden, Daniel L., "Modelling of Choice of Residential Location," in F. Snickers A. Karlquist, L. Lundquist, and J. Weibull, eds., *Spatial Interaction Theory and Residential Location* (Amsterdam: North-Holland, 1978).

McFadden, Daniel L., and Kenneth Train, "Mixed MNL Models for Discrete Response," *Journal of Applied Econometrics* 15 (2000), 447–470.

Nevo, Aviv, "Mergers with Differentiated Products: The Case of the Ready-to-Eat Cereal Industry," *RAND Journal of Economics* 31 (2000), 395–421.

——— "Measuring Market Power in the Ready-to-Eat Cereal Industry," *Econometrica* 69 (2001), 307–342.

Peters, Craig, "Evaluating the Performance of Merger Simulation: Evidence from the U.S. Airline Industry," *Journal of Law and Economics* 49 (2006), 627–649.

Petrin, Amil, "Quantifying the Benefits of New Products: The Case of the Minivan," *Journal of Political Economy* 110 (2002), 705–729.

Reynaert, Mathias, and Frank Verboven, "Improving the Performance of Random Coefficients Demand Models: The Role of Optimal Instruments," *Journal of Econometrics* 179:1 (2014), 83–98.

Small, Kenneth A., "A Discrete Choice Model for Ordered Alternatives," *Econometrica* 55 (1987), 409–424.

Sudhir, K., "Competitive Pricing Behavior in the Auto Market: A Structural Analysis," *Marketing Science* 20 (2001), 42–60.

Train, Kenneth, "Halton Sequences for Mixed Logit," Department of Economics, working paper series 1035, UC Berkeley (May 2000).

Verboven, Frank, "International Price Discrimination in the European Car Market," *RAND Journal of Economics* 27 (1996), 240–268.

Werden, Gregory J., and Luke M. Froeb, "The Effects of Mergers in Differentiated Products Industries: Logit Demand and Merger Policy," *Journal of Law, Economics, and Organization* 10 (1994), 407–426.

Wojcik, Charlotte, "Alternative Models of Demand for Automobiles," *Economics Letters* 68 (2000), 113–118.

APPENDIX

Contraction Mapping for Nested Logit Model

In this appendix we show how to modify BLP's contraction mapping to solve the demand system $s = s(\delta)$ in the random coefficients nested logit model, with a nesting parameter ρ. To simplify the exposition, we consider a nested logit without random coefficients, so $\sigma = 0$. Note that in this case, there is an analytic solution for $s = s(\delta)$ (Berry (1994), so a contraction mapping is not actually needed. The analysis here straightforwardly generalizes to the case where $\sigma \neq 0$.

BLP's Original Contraction Mapping

BLP showed that the function $f(\delta)$, defined pointwise by

$$f_j(\delta) \equiv \delta_j + \ln(s_j) - \ln(s_j(\delta)),$$

is a contraction mapping with modulus less than 1 when the demand system $s = s(\delta)$ is given by the (random coefficients) logit model, where the nesting parameter $\rho = 0$. To satisfy the conditions of their theorem, it is required that f_j is differentiable and satisfies the following monotonicity conditions:

$$\begin{aligned}
&\frac{\partial f_j}{\partial \delta_j} = 1 - \frac{1}{s_j}\frac{\partial s_j}{\partial \delta_j} \geq 0,\\
&\frac{\partial f_j(\delta)}{\partial \delta_k} = -\frac{1}{s_j}\frac{\partial s_j}{\partial \delta_k} \geq 0, k \neq j,\\
&\sum_{k=1}^{J}\frac{\partial f_j}{\partial \delta_k} - 1 = \frac{1}{s_j}\sum_{k=1}^{J}\frac{\partial s_j}{\partial \delta_k} < 0.
\end{aligned} \tag{A1}$$

To assess whether these conditions are satisfied for the nested logit model, note first that the demand derivatives are given by

$$\begin{aligned}
&\frac{\partial s_j}{\partial \delta_j} = \left(\frac{1}{1-\rho} - \frac{\rho}{1-\rho}s_{j|g} - s_j\right)s_j,\\
&\frac{\partial s_j}{\partial \delta_k} = \left(-\frac{\rho}{1-\rho}s_{k|g} - s_k\right)s_j \text{ for } j,k \text{ in the same group } g,\\
&\frac{\partial s_j}{\partial \delta_k} = -s_k s_j \text{ for } j,k \text{ in a different group } g.
\end{aligned}$$

Substituting these derivatives in equation (A1) and rearranging shows that

$$\begin{aligned}
&\frac{\partial f_j}{\partial \delta_j} \geq 0 \Leftrightarrow \rho \leq \frac{s_j}{1 - s_{j|g} + s_j},\\
&\frac{\partial f_j}{\partial \delta_k} \geq 0 \Leftrightarrow \begin{array}{ll} \frac{\rho}{1-\rho}s_{k|g} + s_k \geq 0 & \text{for } j,k \text{ in the same group } g \\ s_k \geq 0 & \text{for } j,k \text{ in a different group } g \end{array},\\
&\sum_{k=1}^{J}\frac{\partial f_j}{\partial \delta_k} - 1 < 0 \ \Leftrightarrow s_0 > 0,
\end{aligned}$$

where $s_0 = 1 - \sum_{k=1}^{J} s_k$ is the market share of the outside good. The second and third inequality are satisfied. However, the first inequality is satisfied only for ρ sufficiently close to 0, so that BLP's function $f(\delta)$ is not necessarily a contraction mapping for the nested logit model.

Modified Contraction Mapping

Consider the following modification of BLP's original function $f(\delta)$,

$$f_j(\delta) \equiv \delta_j + \ln(s_j) - (1-\rho)\ln(s_j(\delta)),$$

which dampens the original function by $(1-\rho)$.

The monotonicity conditions of BLP's theorem become

$$\begin{aligned}
&\frac{\partial f_j}{\partial \delta_j} = 1 - (1-\rho)\frac{1}{s_j}\frac{\partial s_j}{\partial \delta_j} \geq 0,\\
&\frac{\partial f_j(\delta)}{\partial \delta_k} = -(1-\rho)\frac{1}{s_j}\frac{\partial s_j}{\partial \delta_k} \geq 0, k \neq j,\\
&\sum_{k=1}^{J}\frac{\partial f_j}{\partial \delta_k} - 1 = (1-\rho)\frac{1}{s_j}\sum_{k=1}^{J}\frac{\partial s_j}{\partial \delta_k} < 0.
\end{aligned}$$

The second and third conditions remain satisfied as in the orginal contraction mapping for $0 \leq \rho < 1$. To verify the first condition, substitute the above derivative for $\partial s_j/\partial \delta_j$ to obtain

$$\begin{aligned}
\frac{\partial f_j}{\partial \delta_j} &= 1 - (1-\rho)\left(\frac{1}{1-\rho} - \frac{\rho}{1-\rho}s_{j|g} - s_j\right)\\
&= \rho s_{j|g} + (1-\rho)s_j \geq 0,
\end{aligned}$$

for $0 \leq \rho \leq 1$. Hence, the modified function satisfies BLP's monotonicity conditions, and we use this as a contraction mapping to solve the (random coefficients) nested logit demand system. Note that the dampening for the contraction mapping implies a larger value for the Lipschitz constant and thus a slower rate of convergence, especially as ρ approaches 1.